GENERAL LIBRARY
UNIV. OF MICH.
NOV 1 1944

Vol. XXXV SEPTEMBER, 1944 No. 6

The Journal of Educational Psychology

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching

CONTENTS

Rate, Accuracy, and Process in Learning 321
WILLIAM A. BROWNELL

Evolution in Education through Army Experience 338
PAUL A. WITTY AND SAMUEL GOLDBERG

The Relationship between the Centroid and Spearman's Methods . . 347
KARL J. HOLZINGER

The Scoring of Continuity or Rearrangement Tests 352
C. W. ODELL

Relations between Ability and Social Status in a Midwestern Community. I: Ten-year-old Children 357
ROBERT J. HAVIGHURST AND LEOTA LONG JANKE

Difficulty of Vocabulary Learning as Affected by Parts of Speech . . 369
C. L. MORGAN AND D. N. RONHAM

Book Reviews 378

$6.00 per Year • Published Monthly September to May

WARWICK & YORK, INC.
BALTIMORE, MD.

Entered as Second Class Matter Nov. 15, 1921, at the Post Office at Baltimore, Md. under the Act of March 3, 1879; additional entry as Second Class Matter at York, Pa.

THE JOURNAL OF Educational Psychology

Established 1910

EDITED BY

JACK W. DUNLAP
University of Rochester

IN ASSOCIATION WITH

STEPHEN M. COREY
University of Chicago

JOHN G. DARLEY
University of Minnesota

BERTHA PETERSON HARPER
University of Rochester

HAROLD E. JONES
University of California

H. H. REMMERS
Purdue University

PERCIVAL M. SYMONDS
Teachers College (Columbia University)

PAUL A. WITTY
Northwestern University

H. E. BUCHHOLZ
Managing Editor

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY is devoted primarily to the scientific study of problems of learning, teaching, and measurement of the psychological development of the individual. THE JOURNAL will contain articles on the following subjects: the psychology of school subjects; experimental studies of learning; the development of interests, attitudes, and personality, particularly as related to school adjustment; emotion, motivation, and character; mental development and methods. This last will include tests, statistical techniques, and research techniques in cross-sectional and developmental studies.

Dr. Jack W. Dunlap is now on leave of absence from the University of Rochester and serving as Lieut. Commander in the U. S. Navy. For the 'duration', therefore, manuscripts, books and other materials for review, and correspondence regarding editorial matters should be addressed to H. E. Buchholz, Warwick & York, 10 E. Centre St., Baltimore, Md.

Manuscripts should be typed and double-spaced throughout including quotations, footnotes, and references. Return postage should be included with all unsolicited manuscripts.

THE JOURNAL is published monthly from September to May. The price per year in the United States and Pan-American countries is $6.00; $6.20 to Canada; and $6.40 to other foreign countries. Part-year subscriptions are 90 cents per issue ordered. Back volumes are $7.00 each, and back issues $1.10.

Subscribers should notify the Publishers of change of address at least four weeks in advance of publication of issue with which the change is to take effect. Claims for non-receipt of an issue will not be honored unless made within two weeks after receipt of next succeeding number.

WARWICK AND YORK *Publishers* BALTIMORE, MD.

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XXXV September, 1944 Number 6

RATE, ACCURACY, AND PROCESS IN LEARNING

WILLIAM A. BROWNELL
Duke University

GROUP TEST DATA

Group tests consisting of the eighty-one simple multiplication combinations (the 0-combinations were excluded) were administered to about one hundred fifteen children per half-grade, from Grade IIIA through Grade VA, in two North Carolina school systems. The methods of administration and of scoring yielded the usual measures of rate and accuracy of work. Table 1 summarizes very briefly the data obtained.[1]

Fig. 1 presents the data in a somewhat different manner. In the construction of all four curves the records made by the Grade VA pupils were taken as bases for comparison. The median accuracy score for Grade VA was 81; that for Grade IIIA was 64, or seventy-nine per cent of the Grade VA median. The median accuracy score for Grade IVA was 80, or ninety-nine per cent of the Grade VA standard, and so on. The Q_1-curve for accuracy and the median-and-Q_1-curves for rate were constructed in a similar manner, always with the Grade VA measure as the standard.

Inspection of the accuracy curves discloses that progress was most rapid in Grade IVB, when the average child 'knew' the combinations about as well as did the average child in Grade VA, and when three-fourths of the children could write correct answers for ninety-three per cent or more of the eighty-one combinations.

[1] The data reviewed here have been published in much more complete form in the following reference: William A. Brownell and Doris V. Carper. *Learning the Multiplication Combinations*, Chapter IV. Duke University Research Studies in Education, No. 7. Durham, N. C.: Duke University Press, 1943, pp. xii + 177.

TABLE I.—SUMMARY OF DATA FROM GROUP TESTS

Measure	Scores, by Half-Grades				
	IIIA*	IVB	IVA	VB	VA
Accuracy (number correct; 81 possible)					
Actual Median	64	80	81	81	81
Actual Q_1	55	75	80	80	80
Rate (time to finish, in minutes and seconds)					
Actual Median	6:20	4:40	2:45	3:25	2:50
Actual Q_1	8:25	5:30	3:50	3:50	3:25
Actual Q_3	5:20	3:30	2:30	2:45	2:30

* At the time of the testing the Grade IIIA pupils were supposed to have mastered only sixty-five of the eighty-one combinations, the remaining sixteen being taught in Grade IVB.

As judged by rate scores, the greatest gain came in Grade IVA, a half-grade later than in the case of accuracy. In Grade IVB the average rate score was only sixty per cent of what it was in Grade VA; but in Grade IVA, ninety per cent. The Q_1-curve also shows the greatest amount of gain in Grade IVA. As was untrue in the case of accuracy, the children of Grade VB still had some distance to go before they could equal the rate records a half-grade later in Grade VA.[2]

Table 1 and Fig. 1 are typical of reports of research on learning, and the data which they summarize are likewise typical of such reports. They are typical in the sense that they describe two aspects of performance, and only two. Measures of rate and accuracy do, of course, supply information of considerable value with respect to the changes which occur in learning. In

[2] The data for the North Carolina study are in agreement at all substantial points with those obtained in a much more extended study (an average of about six hundred pupils per half-grade) which is reported in Chapter V, Brownell and Carper, *op. cit.* The North Carolina results are dealt with here rather than the results of the larger investigation because of the greater validity of the interview data to be presented shortly.

the present instance, for example, they tell us: (1) the degree to which the experimental children found correct answers for combinations, and (2) the speed with which they wrote their answers. But they tell us no more than this, and more is needed if learning is to be appraised in anything like a comprehensive manner.

As a matter of fact, measures of rate and of accuracy really describe the efficiency of performance. In this study they tell

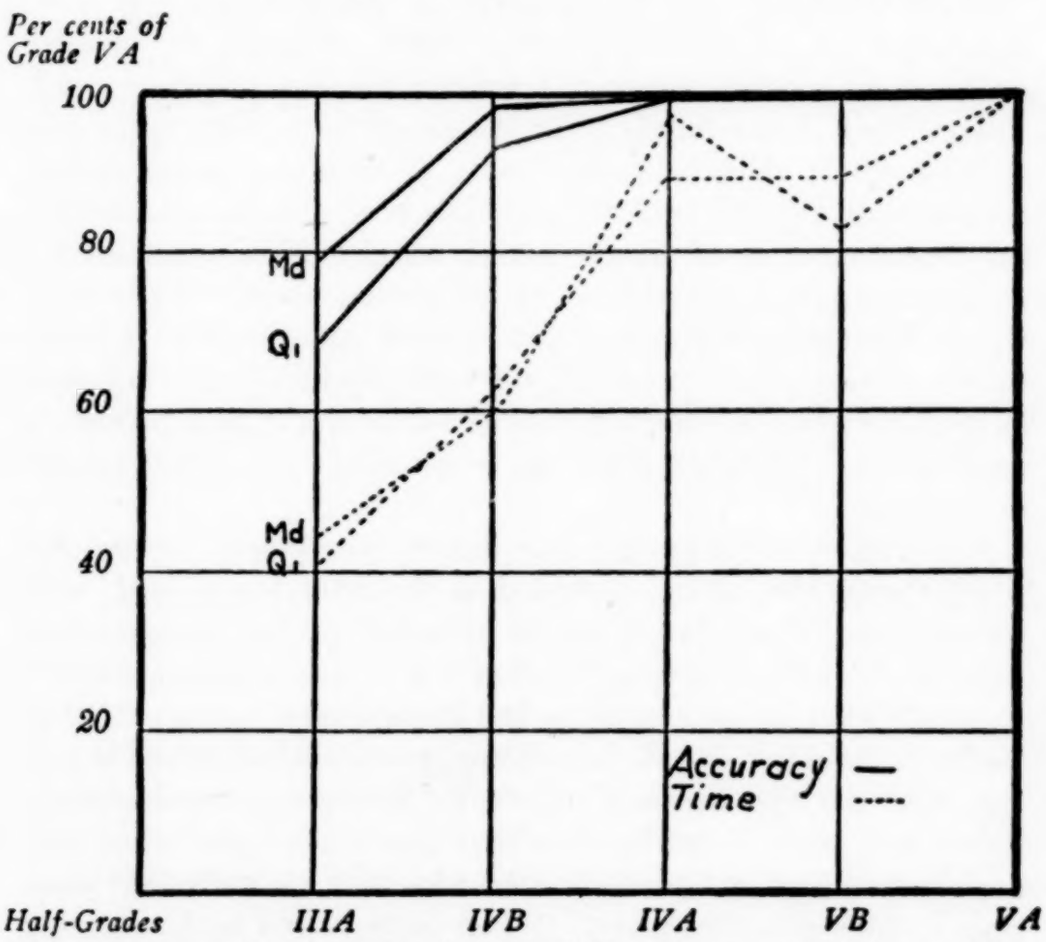

FIG. 1.—Increase in accuracy and rate by half-grades expressed in terms of the Grade VA medians and Q_1's.

us how well (how quickly and accurately) these particular children wrote answers, but they do not tell us how the children got their answers. They depict the efficiency with which processes were used, but have nothing to say about what these processes were. Stated differently, measures of rate and of accuracy do not tell us how good a performance is,—whether it is expert or amateurish, mature or immature, close to the limit of possible improvement or near the first stages of improvement.

To restrict research data on learning to measures of rate and accuracy is to portray learning incompletely. To neglect the processes employed in learning, while it is characteristic of research to do so, is to tell only part of the story,—and possibly the less important part at that.

INTERVIEW DATA

The Procedure.—The group tests on the combinations were supplemented by individual interviews, designed to procure evidence on a third aspect of learning; namely, processes used. The children, immediately after the group tests, were taken one at a time, one child per interviewer. Preliminary research had demonstrated that practically all children's procedures would fall into one or another of ten categories, but provision was made for recording unusual processes. In all grades, except Grade IIIA where the children had not yet studied some of the 'higher' tables, fifteen of the more difficult combinations were employed in the interviews, so selected that all tables would be represented and that no closely related pairs of combinations would appear in succession.

After informal discussion the interviewer placed before the pupil a combination not included in the interview record, with the request, "Can you tell me the answer?" The next questions were of the following types: "How did you get your answer?" . . . . "What did you think to find the answer?" . . . . "What were you thinking after I showed you the combination and before you told me the answer?" Two more practice combinations were then used, to make sure that the child knew what was expected of him and to make sure also that no undesired emotional states were operative. When rapport and understanding had been established, the fifteen interview combinations (nine in the case of Grade IIIA) were presented, and records were kept.

The eleven categories according to which pupils' responses were classified are presented in the chart on page 325.

After a short period of training, the interviewers (all of whom were graduate students in the writer's course, Investigations in Arithmetic) were able to classify children's processes at once and to record their judgments by means of code letters on a mimeographed Interview Blank. Part of this blank is reproduced on page 326. The first two rows carry data for two hypothetical

CHART OF THE ELEVEN CATEGORIES

Code Letter	Description of Class
H	*Habituation, meaningful.*—The correct answer is confidently given at once, with every indication of understanding.
M	*Memory, rote.*—The answer is given at once and with apparent confidence, but there is no evidence of understanding.
G	*Guessing.*—The answer, usually wrong, is given rather promptly, but it is evident that the child is guessing.
S	*Solution.*—The child starts with a familiar combination and adds to, or subtracts from, the product to get the answer. There are many kinds of solution, for example: $3 \times 4 = 2 \times 4 = 8, +4 = 12$ $3 \times 4 = 4 \times 4 = 16, -4 = 12.$
C	*Counting.*—The child adds or counts the same unit several times. $3 \times 4 = 3, 6, 9, 12$; or, 4, 8, 12.
R	*Reversal.*—The child interchanges multiplier and multiplicand to get a more familiar order. $6 \times 3 = ?; 3 \times 6 = 18$, so $6 \times 3 = 18$.
T	*Tables.*—The child starts with a lower combination in the table involved and recites the table to the required point. $4 \times 6 = ?; 4 \times 1 = 4; 4 \times 2 = 8; \ldots 4 \times 6 = 24$ or, $4 \times 4 = 16; 4 \times 5 = 20; 4 \times 6 = 24.$
V	*Visualization.*—The child reproduces groups of objects or figures in clear imagery and works with these. $4 \times 5 = ?; 4 \times 5¢$ (nickel) $= 20¢$, or 20.
N	*No attempt.*—The child makes no attempt to state the answer.
I	*Indeterminate.*—The child gives an answer for the combination, but you cannot ascertain how he got it (that is, his process).
X	*Miscellaneous.*—The child reports a process not included in the preceding list—or you are uncertain as to its proper classification. Report what the child said, using the back of the Interview Blank if this is necessary.

children, to serve as models for entering answers and processes. And it should be added that the names and code letters of the eleven categories were included at the bottom of the blank for easy reference.

In general, the eleven categories previously mentioned were found to be serviceable and reliable. To begin with those at the bottom of the list, in only one half-grade (IVB) were as many as 1.2 per cent of the responses classified as X (miscellaneous); in Grade IIIA the per cent was 0.7, and in the other

INTERVIEW BLANK

School_____ Town_____ Date_____
Teacher_____ Grade_____ State_____ Interviewer_____

Pupil's name	7/5	6/9	3/8	6/5	8/7	6/4	3/7	9/8	7/6	9/2	4/7	5/3	8/6	7/9	3/6
Joe H.	35 H	54 N	24 R	30 H	56 T	24 H	21 H	72 X 65 + 8	42 T	18 H	28 H	15 H	48 T	52 T	18 H
Susie Q.	35 H	54 S	24 R	30 H	56 S	24 H	21 H	75 G	42 T	18 H	28 H	15 H	48 S	60 S	18 H

three half-grades, 0.1 or less. In only one half-grade (VB) were processes finally classified as I. Of course at the time of the interviews more procedures than these were tentatively identified as indeterminate; but because of the complete records made then, it was possible later to dispose of most of them in one or another of the first nine categories. The fact that in only one half-grade (IVB) I and X together contributed more than one per cent of the classifications may be taken as evidence both of the degree to which the nine other categories were adequately comprehensive and of the degree to which identification in these nine categories was possible.

To continue to read the list of categories upward, it may be said that no difficulty arose in recognizing when children visualized (V), or recited tables (T), or reversed combinations (R), or counted (C), or solved unknown combinations from better known combinations (S). It was, however, not so simple to distinguish

among guessing (G), the use of meaningless associations (M), and the use of meaningfully habituated associations (H). However, as will be explained in a later section, means were found to meet this problem, and it is believed that the data for these three categories as they are here reported are about as reliable and as valid as are the data for the other categories.[3]

The data obtained.—Table 2 for Grade IVB is typical of the tables which summarize the interview data for the five half-grades (except that in Grade IIIA the interview combinations included only the nine which had been taught). In Grade IVB fewer than half of the combinations (2040 in all,—one hundred thirty-six subjects, fifteen combinations each) were dealt with by processes which could be classified as H. Three other types of procedure, M, S, and R, together contributed about a third of the total, each of the three to about the same extent. Guessing occurred with about six per cent of the combinations, and counting with about four per cent. Answers were not attempted in forty-seven instances (two per cent) and were obtained by some form of counting in fifty-three instances (three per cent).

Theoretically, at least, the teachers of these children had been striving for a year and more to develop the kind of thought process here designated as meaningful habituation (H). True, unwittingly they had encouraged M-responses, for their methods of teaching from the outset stressed repetitive practice on the formulas of the combinations. Still, it is surprising to discover how far these children were in Grade IVB from the expected kind of procedure. In spite of considerable drill (and, correspondingly, in spite of the frequent repetition of the number facts), these children had not yet abandoned immature and, in certain cases, undesirable procedures (M and G, notably). Indeed, in view of the method of instruction one may well ask how these children happened at all to be using other procedures

[3] This statement cannot be made for the extended study—Brownell and Carper, *op. cit.* The latter study drew subjects from seven other states than North Carolina and also from schools in North Carolina not easily accessible to the writer and his students. On this account in the extended study: (1) directions for the interview had to be mailed to the officials and teachers who did the interviewing in each school, and (2) it was impossible to supervise the actual interviewing as was done in the local study. As has been stated, the group test data in the two studies are closely comparable, as are also the interview data for the lowest three half-grades.

than M, N, and G (and possibly H), for the conditions of learning were not calculated to suggest other procedures.

Fig. 2 is based upon the five tables of interview data of which Table 2 is a sample. It pictures learning in a different way than does Fig. 1, where the criteria of learning are rate and accuracy. In Fig. 2 the criterion is nature of the processes employed to find

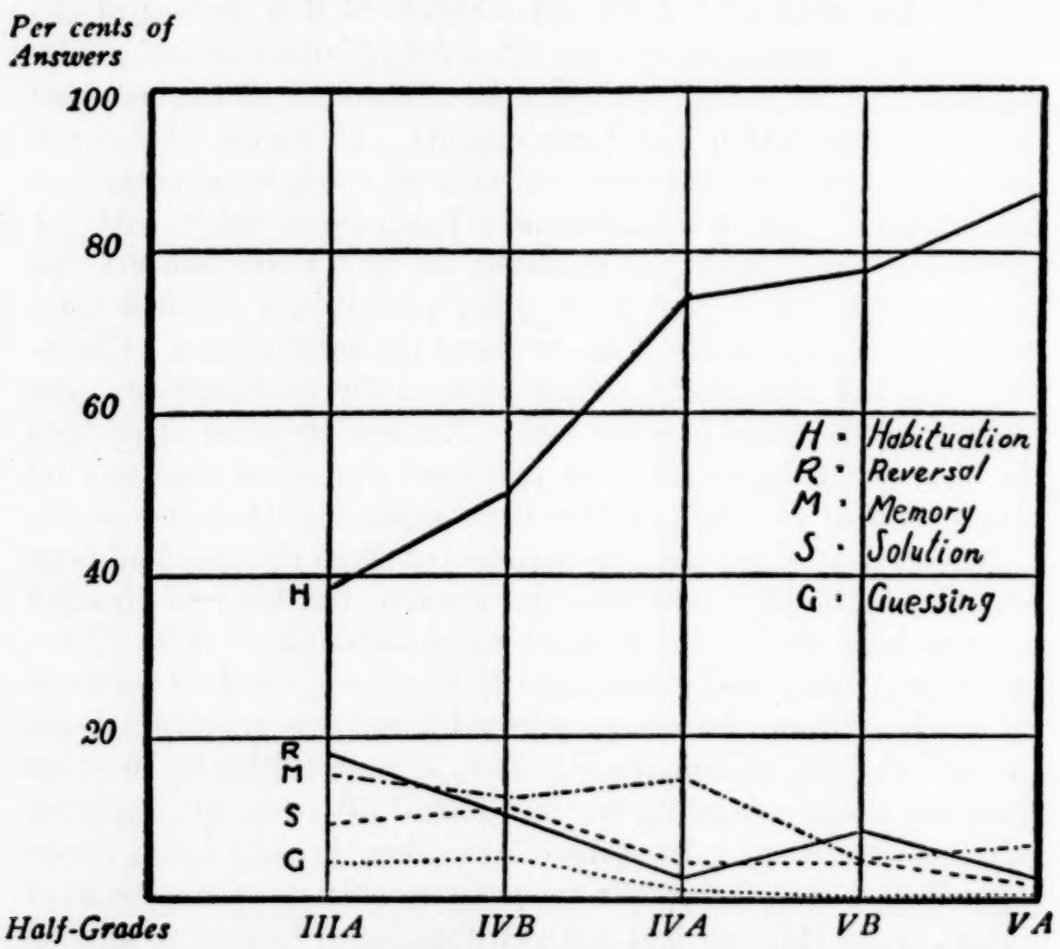

FIG. 2.—Changes in processes used from grade to grade, as revealed in interviews.

answers. The curves for the five major processes represented in the figure need to be studied with due regard for their relationships. As compared with Grade IIIA, there is some gain in the use of H in Grade IVB, with slight compensating losses in use of R and M. The marked gain in use of H in Grade IVA is made possible by the practical abandonment of G and by the tendency to discard R and S. The most striking change in Grade VB is a reduction in the use of M, there being little increase in use of H. The situation in Grade VA is that about twelve

per cent of the combinations were still obtained by methods short of H,—which is to say that even after two and a half years of instruction these children had not developed the most desirable and the most mature procedure for all combinations.

TABLE II.—GRADE IVB: PROCESSES USED WITH SELECTED COMBINATIONS AS REVEALED IN INTERVIEWS (N = 136)

Process	7/5	6/9	3/8	6/5	8/7	6/4	3/7	9/8	7/6	9/2	4/7	5/3	8/6	7/9	3/6	Total
H	79	56	47	78	52	79	60	67	68	93	55	91	63	46	57	991
M	20	18	15	16	17	15	16	24	13	18	16	18	17	27	19	269
G	7	10	6	4	21	4	2	12	14	2	7	1	13	14	2	119
S	11	10	17	22	14	18	15	11	20	6	19	14	15	19	18	229
C	6	4	4	3	5	7	4	6	4	10	4	6	11	4	6	84
R	5	24	41	5	9	5	37	3	4	4	26	2	3	12	30	210
T	5	4	5	2	8	6	2	1	4	3	6	2	2	..	3	53
V	1	1	..	1	..	1	..	1	..	..	..	1	2	1	1	10
N	1	6	..	2	8	..	..	8	5	..	..	..	7	10	..	47
I	..	1	..	..	..	..	..	1	1	..	..	..	..	..	..	3
X	1	2	1	3	2	1	..	2	3	..	3	1	3	3	..	25

Relation of Interview Data to Group Test Data.—The facts revealed by Fig. 2 should now be placed alongside the facts pictured in Fig. 1, for they serve to explain the findings with respect to gains in rate and accuracy. The Grade IIIA children were slow and inaccurate on the group tests because they had meaningful habituations for fewer than half of the (interview) combinations and because they had to employ with combinations which they had not yet habituated meaningfully, processes far short of H in maturity and efficiency. These processes, especially guessing, were susceptible to error, and they were employed clumsily and none too successfully, with the consequence of increased time and mistakes.

The rate curves (Fig. 1) show moderate increases, and the accuracy curves considerable increases in Grade IVB. Improvement in these aspects of learning is explained, less by the abandonment of immature and inefficient processes than by greater control over the same processes which characterized the children

in Grade IIIA. That is to say, the children in Grade IVB used the processes of Grade IIIA, but used them more effectively. Use of H increased only from thirty-nine to forty-nine per cent.

In Grade IVA the slight gains in accuracy brought the experimental subjects practically to the point of proficiency evidenced in Grade VA. The gains in rate are very marked. Again, the interview data provide the reasons. Three-fourths of the combinations were now obtained by the use of H, one-fourth more than in Grade IVB. Only one answer in ten was secured by immature (but desirable) processes, and only one in six (compared with one in five in Grade IVB) by undesirable processes.

The group test data for Grade VB (Fig. 1) reveal no changes in accuracy as compared with Grade IVA results. The maintenance of the Grade IVA level of accuracy is consistent with the lack of change in processes (Fig. 2). The temporary loss in rate is explained by the fact that for some time prior to the group testing these children had had little practice on the multiplication combinations.

The Grade VA children had practically ridded themselves of immature processes and of undesirable processes on the interview combinations (and supposedly on the other combinations of the group tests as well). Use of H had increased to eighty-eight per cent. These facts together account for the excellent showing in rate and accuracy, but they also disclose that, so far as process is concerned, these children were still short of final maturity of method (H).

SIGNIFICANCE OF PROCESS IN LEARNING

Implicit in the foregoing discussion is the notion that maturity of process is a better index of learning than are the customary indices of rate and accuracy.[4] Certain processes have been

[4] Caution requires that this statement be restricted to the kind of learning investigated in this study; namely, the acquisition of meaningful relationships. It is conceivable that level of process is not so good a criterion in the case of other kinds of learning, in which meaning is not an important aspect. But the multiplication combinations are properly learned only when they are meaningfully learned: they must be retained and used in countless differing situations. Perhaps it is not unfair to say that the learning task in the case of the multiplication combinations is more typical of those which are set (or should be set) in the classroom, than are the meaningless learning tasks commonly set in research on learning.

described as 'mature' and 'immature,' as 'desirable' and 'undesirable,' as 'effective' and 'ineffective.' For example, G is undesirable, immature, and inefficient (unless by some chance the correct answer is guessed). R is fairly effective but immature (since it is short of H), but not undesirable as an early way of thinking of combinations. M is undesirable (since the formulas may be learned without understanding) and effective only when the correct answer has been memorized and when the given combination is used in situations which to the learner are very similar to the learning situation. S, like R, is immature but relatively effective; and it is desirable in the early and intermediate stages in learning, for R involves the apprehension of relationships and so contributes to understanding. C is immature, none too effective (especially when a good deal of counting is needed to secure the answer), and is desirable only at the outset of learning. T occupies about the same place as C in the scale of maturity, effectiveness, and desirability, except that it is more likely to be associated with memorization. At the bottom of the scale of maturity is G; at the top, H.

Now, there is no necessary relationship between level of process used and scores in rate and accuracy. A process low in maturity (e.g., C) may be so completely mastered as to make both for speedy and for correct responses. On the other hand, a process higher in maturity (e.g., S) may be so imperfectly mastered that errors and excessive time result.[5] It is for these reasons that at an earlier place in this paper the popular measures of rate and accuracy were termed measures of efficiency: they reveal only how well (promptly and correctly) a given child uses whatever method he does use. It is for these reasons also that the maturity of processes used has been suggested as a sound index of learning. And it is for these reasons, again, that Fig. 2 is offered as a complement to the traditional so-called learning curves which are constructed from measures of rate or of accuracy alone.

---

[5] One consequence of this lack of relationship between level of process and efficiency of performance is that test scores, usually expressed in terms of rate or of accuracy, are frequently misleading. High scores in rate or in accuracy, or in both, are commonly accepted as evidence of sound learning. Such may not be the case at all.

Learning as re-organization.—The impression one gets from typical learning curves is that improvement consists simply in the refinement and the mastery of the response with which one starts initially: one practices this response over and over again, and the result is a smoother, quicker, and more accurate performance. The conception of learning which is revealed by Fig. 2 is something quite different: learning consists in the progressive re-organization of processes or procedures. At the outset the learner satisfies the requirements of the task by using relatively cumbersome and ineffective procedures. He practices at this level for a long or short period, thereby getting greater command over (efficiency with) this type of performance. Continued practice will not raise him to the next higher level of process. Instead he must, for one reason or another, re-organize his response in order to move on the next higher level. Here again he may pause to secure more or less mastery of that process, —only again under appropriate motivation, to abandon this level for the next above. Eventually, after a series of re-organizations, he may arrive at an expert, adult-like level of performance.

TABLE III.—COMPARISON OF PROCESSES USED, FALL AND SPRING, WITH THE SAME SEVENTEEN SUBJECTS, OBTAINED BY INTERVIEWS

Fall Processes, with Frequency	Spring Processes, Same Combinations: Frequency						
	G	M	C	T	R	S	H
H(18)	..	1	..	..	1	..	16
S (49)	..	3	..	5	..	7	34
R (10)	..	3	..	..	3	1	3
T (21)	1	4	..	1	4	1	10
C (43)	1	1	1	3	..	4	33
M (58)	..	35	..	..	1	..	22
G (25)	..	11	..	..	2	4	8
N (18)	..	1	2	..	1	7	7
V (3)	..	..	..	..	..	..	3
X (10)	2	4	..	..	1	2	1
Total (255)	4	63	3	9	13	26	137

This conception of learning is to be seen in Fig. 2 for the group of experimental subjects as a whole. It is to be seen too in Table 3, in the case of seventeen pupils who were interviewed in the fall when they were in Grade IVB and again in the spring, six months later, when they were in Grade IVA.

In Table 3 ten categories of process are listed in the first column in the order of their approximate maturity, with H at the top and C, N, V, and X at the bottom. They are listed again (seven, this time) along the tops of the next columns, with the least mature processes at the left. Of the eighteen combinations dealt with by H in the fall sixteen were answered in the same way in the spring, one process being then classified as M, and another as R. S, a comparatively mature way of thinking of the combinations, provided a sound basis later for H: thirty-four of the forty-nine combinations solved in the fall had been meaningfully habituated by spring. Four of the ten combinations answered by R in the fall were dealt with by higher processes (S and H) in the spring. The movement toward more mature procedures is shown also in the case of the use of T and C. On the other hand, the slight contributions of M and G to meaningful habituation are equally as evident in the figures of Table 3. G and M could hardly evolve into H because they themselves are devoid of meaning.

While the interview data warrant the conception of learning as a series of progressive re-organizations, one must not go too far in assuming greater uniformity than actually exists. One should not infer, for example, that all children in a given grade are at the same stage of maturity of process with respect to all combinations. The figures of Table 2 discredit any such notion. Nor, again, should one infer that any given child is at the same stage with respect to all combinations. Data on this point, which show the variety of processes which a group of children employed with different combinations, are contained in Table 4, which provided the basis for Table 3. And still, again, one should not infer that all children traverse the same series of stages preliminary to final meaningful habituation. On the contrary, beyond any question some children omit one or two types of process (they may, for example, not use R or T with some combinations), and some move toward H much more rapidly than do others. Still, with all these limitations, the general

picture of progressive reorganization as revealed in Fig. 2 seems to be sound. The learning of meaningful relationships is no simple straight-line affair. Instead, such learning is characterized by a continuous series of changes from lower or less mature and effective organizations of behavior to steadily higher or more mature and effective types of response.

DISCUSSION

The study of the learning of the multiplication combinations is reviewed here less because of its practical than because of its theoretical implications. Research on learning rarely reports (and probably rarely secures) measures of the kind collected in the present instance by means of the interview. As a consequence, the accounts of learning which emerge from such research must necessarily be incomplete, and they may well be misleading. The nature of the incompleteness and of misinterpretation has been illustrated in the foregoing pages.

It seems curious that learning in the case of meaningful tasks should be so commonly investigated with no regard for what happens with regard to the meanings themselves. Perhaps one reason for neglect of data at this point is that the meaning-aspects of behavior are difficult to quantify. In turn, they are difficult to quantify because meanings are peculiarly subjective. Hence, it is seldom possible through objective tests to secure trustworthy measures. This obstacle will, however, disappear as soon as we are willing to substitute other techniques for that of objective testing,—as soon as we are willing, for example, to employ such evaluation procedures as those of the interview and of direct observation.

There is sound justification for the suspicion which attaches to these procedures. Unless they are carefully planned and managed, observation and the interview are subject to all sorts of error. This fact,—and it is a fact,—merely signifies that proper safeguards must be found and applied. And they can be found and applied.

Attention has already been called in the section, Interview Data, to the difficulty of discriminating among H, M, and G when children quickly gave their answers, correct or incorrect, to number combinations. On this account it became necessary to modify the interviewing technique. Changes were made in

TABLE IV.—PROCESSES USED BY SEVENTEEN GRADE IVB PUPILS ON THE INTERVIEW COMBINATIONS WHEN PRESENTED TWICE, IN FALL AND SPRING

Combination	Subjects, with Processes Used in Fall and Spring																
	HK	MS	EM	RH	MM	JM	CD	BR	GJ	SS	AM	CG	MR	RF	DH	AH	BG
7	C*	M	G	T	M	T	S	G	M	C	V	M	H	G	S	T	M
5	H	M	M	G	M	H	S	H	H	H	H	H	H	H	S	H	M
6	C	M	M	X	M	N	S	N	M	C	N	M	S	C	S	T	R
9	H	M	M	G	H	S	S	H	H	H	H	M	H	H	T	R	M
3	C	M	G	X	M	S	S	R	M	S	R	R	C	C	C	T	G
8	H	M	M	G	M	H	H	M	H	H	H	H	H	S	T	R	R
6	C	M	N	X	M	T	S	G	M	C	V	S	H	C	S	T	T
5	H	M	M	M	H	H	S	S	H	H	H	M	H	H	H	H	M
8	C	M	G	X	M	N	T	N	G	C	N	G	S	C	S	G	M
7	H	M	M	S	M	S	M	S	S	G	H	M	S	S	T	S	R
6	C	M	G	X	M	S	S	G	M	C	T	S	S	S	S	T	N
4	H	M	M	G	M	H	H	H	H	H	H	H	H	H	T	H	S
3	C	M	M	H	M	S	S	C	M	S	T	S	H	C	C	R	R
7	H	M	M	M	H	H	H	H	H	H	H	M	H	H	T	R	R
9	C	M	M	X	M	N	M	G	G	C	N	M	H	G	G	M	R
8	H	M	M	M	M	H	H	S	H	H	H	M	H	H	H	M	M
7	C	M	G	X	M	S	S	N	M	S	N	M	S	G	C	T	G
6	H	M	M	S	M	S	H	S	H	H	S	H	H	H	T	M	R
9	C	M	M	M	M	S	C	C	M	H	M	H	H	H	H	M	C
2	H	M	M	M	H	H	H	H	H	H	H	H	H	H	H	H	M
4	C	M	G	X	M	T	S	S	M	S	S	C	S	C	S	T	N
7	H	M	M	M	H	H	H	H	H	H	S	H	H	C	T	R	R
5	H	M	G	H	M	S	S	C	H	H	V	S	H	H	C	T	M
3	H	M	M	R	H	H	H	S	H	H	H	H	H	H	H	H	M
8	C	M	G	X	M	N	S	N	R	C	N	S	S	C	G	T	M
6	H	M	M	M	M	C	H	H	H	H	H	H	H	H	H	S	M
7	C	M	G	M	M	T	S	N	S	C	N	S	S	C	S	T	R
9	H	M	M	M	M	T	H	G	H	H	S	M	H	H	T	M	R
3	C	M	M	M	M	S	S	C	M	S	T	S	H	H	C	T	G
6	H	M	M	H	H	H	H	S	H	H	H	H	H	H	H	R	M

* The upper letter in each pair refers to the process used in the fall; the lower letter, to the process used in the spring.

the following manner: When children gave answers, correct or incorrect, which they could not justify, instead immediately offering other answers when questioned about their first attempts, it was assumed that they had guessed, and G was recorded as the process. In the case of H and M, (1) a tentative entry of M or H was recorded at once when one of these processes seemed to have been used. At this time questioning was carried only far enough to give a reasonable basis for the preliminary classification. Then, after the subject had responded to all fifteen combinations, he was taken back over the M and H facts. (2) On the second trial the child was asked, "How did you know that . . . was the correct answer?" and "What did you think to get the answer?" If these questions did not elicit an unambiguous report, an absurd suggestion was offered. For example, for 8×3 it might be asked, "Did you get your answer by thinking '8 and 3 are 11, and 7 are 18, and 6 are 24'?" The child, in rejecting this absurd procedure, usually explained what he had done. If so, the appropriate entry was made, and the next doubtful combination was presented. If not, the original entry of M or H was allowed to stand until the remaining doubtful facts had been considered. The child was then taken back over the uncertain M and H facts. (3) On the third trial the interviewer said, "But suppose I say that 24 (for 8×3) is not correct. Suppose I say the answer is 25. How could you prove me wrong?" If the child was able by any rational and logical method to prove the correctness of his answer and the incorrectness of the suggested answer, it was assumed that he understood what he was doing, and he was given credit for H. If, on the other hand, he had no way of testing the correctness of answers, his process was recorded as M.

This whole procedure sounds much more complicated in the description than it was in actual use. The children were not aware of the three levels of questioning. Their responses were not distorted by the inquiry, but on the contrary remained true to reality. Indeed, only rarely did children realize that they were asked to deal as many as three times with the same combinations.

The modified technique employed in interviewing is described here in detail in order to demonstrate the feasibility of surrounding this research procedure with safeguards which will insure trustworthiness of data. The peculiar advantages of the inter-

view are not infrequently lost through failure to recognize the dangers which it entails. In research such as that here reported the interview cannot be completely standardized. Instead, the questioning, which of course must follow a general pattern, must be left flexible enough to elicit responses even from uncoöperative children. Furthermore, it must be flexible enough to allow for individual differences in understanding the situation to which they are to react. But this flexibility must be kept within limits. Otherwise, the interviewer who has no clear insight into his problem can readily ruin his data by suggesting to his subjects processes which, while they are not their own, they will nevertheless claim as their own.

So much for the interview as a research technique in learning research: It is exceedingly valuable if it is sagaciously employed. It is valuable because it enables the investigator to secure a kind of data which are usually not collected on the ground that they are likely to be unreliable. But the data which the interview can supply are essential to a complete account of the learning process. They must therefore be secured. Otherwise, we shall continue to over-simplify and to misunderstand the true course of meaningful learning and hence to base instruction upon an imperfect foundation.

EVOLUTION IN EDUCATION THROUGH ARMY EXPERIENCE

PAUL A. WITTY
Major, A.G.D.
AND
SAMUEL GOLDBERG
Captain, A.G.D.

In a recent article, Major General Edmund B. Gregory, the Quartermaster General, U. S. Army, writes: "You may not know it yet, but what your Army's Quartermaster Corps is doing now will revolutionize your life after the war. It will affect your clothing, your food, your house, even your recreation." Convincingly, he describes new wear-resistant clothes; new sterilizing rinses; a toughened type of glass; new soaps; developments in prefabrication and cheaply installed central heating systems; developments in dehydration, quick-freezing and concentrates; and dozens of other innovations which will make our efforts more efficient and our daily life more pleasant. The Quartermaster Corps in helping our soldiers become more efficient fighters has succeeded also in enriching daily living.

Similarly practices in the Army's training program will greatly influence civilian education. These practices will be described and programs in which they are followed will be cited. Finally, educational principles which the Army has followed successfully will be set forth. They should receive serious attention in educational planning.

Training is a serious affair in the Army. It must be accomplished with maximum efficiency. To keep instructors ever conscious of the importance of their mission, posters containing the following admonition are exhibited widely in posts, camps, and stations: "Let no man's ghost damn your training program." This script is effectively super-imposed on a picture of a grave with a helmet over the cross.

An integral part of all training is the proper education of both officers and men. Due appreciation of the rôle of education in the expeditious training of the American Army was expressed early in the mobilization program by General Somervell, Commanding General, Army Services Forces, when he stated: "We

can lose this total war on the battlefront as a direct result of losing it on the educational front. Education is the backbone of an Army." Accordingly, some of the best educational practices and procedures were introduced and used in the mobilization training program. Men were provided with training which would lead them to develop the highest degree of proficiency in the shortest amount of time. It was necessary to insure success in combat. It was essential to achieve this objective with great speed. There was little time to lose when America entered the world struggle as an active belligerent.

PRE-INDUCTION TRAINING AND THE ARMY SPECIALIZED TRAINING PROGRAM

The Pre-induction and Army Specialized Training Programs are indications of the completeness with which the Army has approached its manpower needs. Technology and logistics are important in this mechanized, mobile war. But this global conflict will also require millions of men before the forces of barbarism and tyranny are defeated. It takes a short time now to turn out a plane or a tank, but it takes many months to produce a pilot or a tankman. To expedite the training of selectees, the Army initiated a Pre-induction Training Program. Pre-induction training of secondary-school youth equips many sixteen- and seventeen-year-olds for the day when they will enter the armed forces.

In response to the Army's needs, many secondary schools have provided students with (1) an understanding of the issues of the war, (2) an appreciation of the nature of Army life, and (3) a command of basic language and computation skills. Physical hygiene programs have been intensified to produce high-school graduates who are physically fit. And, many students have been given training along mechanical, manipulative, and technical lines so that they may serve in "one of the Army's more than 650 occupations."

To achieve the goals of pre-induction training, many secondary schools have modified their curricula and methodology. Others have intensified some of their more desirable practices. The demands of pre-induction training have made it necessary for secondary schools to deal realistically with the issues of the war. General Dalton, Director of Personnel, A.S.F., expressed

this need in a recent article: "The sixteen- and seventeen-year-old boy must know how this war came about, who our enemies are, who is fighting on our side, what it would mean to our way of life if we lost, what it can mean to humanity if we win." Students are receiving orientation courses which are truly deserving of the title, "Social Studies." They are developing a morale which grows from an understanding of the basic issues of this struggle.

When juniors and seniors in high schools have become familiar with Army procedures in Induction Stations, Reception Centers, and Replacement Training Centers, and with Army organization and life, generally, they will find it easier to make the shift from civilian to military status. They will be less overwhelmed by the demands of Army discipline and regulations than were the earlier selectees. Their adjustment will be less complicated, their frustrations and irritations will be fewer. Their familiarity with the Army setting will give them a feeling of confidence and security. They will make happier and more effective fighting men.

Physical hygiene courses are stressing first aid and conditioning activities. Proper emphasis is now being given to those basic skills of communication and computation which are essential in our technological culture and society. And, more students than ever before are receiving technical, manual, and scientific courses.

It is apparent that under the stress of Army influences, secondary schools are learning to provide a type of education more clearly associated with students' needs and interests. Some are developing a kind of polytechnic orientation. It is well for them to get away from the idea of academic, general, commercial, manual, vocational, and technical programs as such—with each containing its separate departmentalized courses. Fundamentally, secondary schools should provide all students with adequate language, number, and technical skills, and with an understanding of human relations in group living. All adolescents need such orientation and education.

It is possible now to discern how undesirable it was to provide technical and scientific training for vocational, manual, and technical high-school students and not to include such experience for academic, general, and commercial students; to offer courses in essential commercial arithmetic and related fields to commercial students and make little provision for such courses in the aca-

demic, general, and technical courses; and to give courses in language and social studies in abundance to general and academic students and to allow only fragmentary contact with these fields for commercial, vocational and technical students.

Specialized Training is another program through which the Army has influenced educational practice. The Specialized Training Program was organized to supply the needs of the Army for men who are trained in medicine, engineering, languages, science, mathematics, and psychology. It became necessary for the Army "to assure itself that there would be no interruption in the flow of professionally and technically trained men," when the minimum selective service age was set at eighteen.

Through the Army Specialized Training Program, the facilities and faculties of selected colleges and universities were made available to qualified soldiers. These men were given training in courses prescribed by the Army. The curricula which were developed were highly practical in nature since the men were to be taught skills they would need in the Army. There was a minimum of overlap in the content of different subjects. It was requested of the teaching faculties that (1) abundant use be made of training aids, (2) demonstration exercises be organized, (3) reviews be frequent, (4) testing be periodic, and (5) ample opportunity be provided for applying acquired information and skills. Curriculum and method were developed and coördinated to insure high motivation, rapid assimilation of subject-matter, and the acquisition, retention, and application of skills.

Classification personnel were assigned by the Army to the various colleges and universities to aid in the individual study of men who were experiencing scholastic and personal difficulty. The classification sections worked closely with the academic faculty, military instructors, and the student soldiers. Accordingly, there was a better coördination of the man's training. Moreover, his progress was systematically recorded.

As a result of participating in the Army Specialized Training Program, higher education cannot help but be more critical of some of its past practices. Many subjects have been included in college curricula with insufficient thought and attention concerning the objectives they were to accomplish; various subjects overlapped to a considerable degree; too heavy a reliance was placed on the lecture method; inadequate use was made of visual

aids; and too few opportunities were provided in many areas for applications to be made.

The effects of the Army Specialized Training Program will be felt not only in those areas of higher education concerned with curricula and methodology, but also in the area of guidance. Certain values of a sound personnel system have been demonstrated.

CHARACTERISTICS OF ARMY TRAINING

It is not feasible to describe all phases of basic training and the different varieties of technical and specialist training which the Army program encompasses. However, a brief description of two different types of training given in the Army should serve to illustrate important educational principles.

Nine out of every ten soldiers must be trained for specialized jobs. This is true of men who engage in actual combat and those who perform service functions. For example, more that 150 Army jobs require a knowledge of electricity; upwards of 200 necessitate a basic knowledge of machines; and more than 175, a knowledge of shopwork. To fulfill these requirements, it is often necessary to train many inducted soldiers for assignment to jobs which are unrelated to their civilian experience.

Assignment is not and cannot afford to be haphazard. Each man is examined at a Reception Center to determine his general learning ability and his special aptitudes for mechanical and radio-telegrapher occupations. Then, he is interviewed concerning his civilian occupational experience. Other abilities which may have been acquired through avocational pursuits are also ascertained. On the basis of interviews and the test data, each man is designated for a given branch of the service. He is sent to an appropriate Replacement Training Center or organization. At this stage, he is again interviewed. If necessary, he is given further aptitude tests and examinations. Finally, he is recommended for specific job training within a particular branch of the service.

After the man has completed his basic field training, he is given technical training. Let us assume the man is to be trained as an automobile mechanic in some branch of the service. He attends a mechanic's school in which the finest equipment is found. The instructor personnel is carefully selected and thoroughly trained. The ratio of instructors to trainees is approximately one to

thirteen. The relationship between the instructor and the student is a warm personal one, since each realizes the seriousness of his job and the importance of his mission. Lessons are well graded, and students who have special difficulty are given individual help. Instruction is mainly of a practical nature. Testing is an integral part of teaching and is used for diagnosis and evaluation as well as for rating.

Despite careful procedures of selection and well directed classroom instruction, some soldiers do not succeed in the training. In such instances, they are re-interviewed and reclassified. Every effort is made to fit them into a training program in which they will succeed.

The training program for illiterate, non-English speaking and educationally retarded men also serves to illustrate educational practices and procedures applied in the Army.[1] Not all illiterates and non-English speaking men are accepted for special training. Through the use of mental tests in the induction station, many illiterates and non-English speaking men who are probably feeble-minded are screened out and rejected for military service. Those whose illiteracy and inability to use and understand English seem to be the result of environmental circumstances and not of low mental ability are sent to special training units.

In special training units, the men are examined further with placement tests, designed to measure their degree of functional literacy. Test results are used to place them in proper levels. They are then assigned to classes, the average size of which is fifteen. The small size of classes makes it possible for the teacher to differentiate instruction.

Instructional materials have been prepared which are highly functional in character. They are carefully graded and are related to the lives of the men in the Army. The materials are designed so as to make it relatively easy for the teacher to integrate his instruction in the various fields, i.e. reading, writing, and arithmetic.[2]

---

[1] Paul A. Witty and Samuel Goldberg, "The Army's Training Program for Illiterate, Non-English Speaking, and Educationally Retarded Men," *Elementary English Review*, XX (1943), 306–311.

[2] Paul A. Witty and Samuel Goldberg, "The Use of Visual Aids in Special Training Units in The Army," *Journal of Educational Psychology*, February, 1944, 82–90.

Tests have been prepared which instructors can use to measure the accomplishment of their men. It is possible to use these examinations analytically to obtain a diagnostic inventory of individual and class difficulties. Special publications have been prepared which describe various teaching procedures that can be used profitably in the classroom. The development of training aids for academic and military instruction has been encouraged and directed. Training aids sections can be found in most special training units and in most training organizations throughout the Army.

The approach to each man is positive and constructive. A cumulative record is kept of his progress throughout training. When a man has attained desired academic and military standards in a special training unit, he is sent to an organization for regular basic training. If, at the end of the period of special training, the man is unable to achieve the indicated levels of accomplishment, he is then considered for honorable discharge from the Army.

Teachers are carefully selected for work in special training units. National training conferences are held to acquaint the instructor personnel with approved teaching methods. Representatives from the various units who attend the national conferences then initiate local conferences in their units and service commands to give their colleagues the benefit of their newly acquired information. Troop schools are conducted by most units to insure that the training staff receives continual indoctrination in the best methods of instruction. Supervisory 'inspections' are made periodically to make certain that each unit is maintaining a satisfactory level of training.

The program of special training has proved to be successful in the Army. It has provided many leads for the development of a comparable civilian program of adult education.

CONCLUSIONS

Experiences gained in various phases of the Army's training program have served to demonstrate the value of many educational principles. Before giving consideration to them, it is important, however, to observe certain cautions. The training of men in the Army does not present the same problems as the education of individuals in elementary and secondary schools and

colleges. Problems of motivation are different. The experiential background of the adult student is radically different from that of students at the elementary, secondary, and college levels. Finally, though courses of study can be quite definitive in the Army, it is not possible to be so precise and certain about school curricula and objectives.

Despite apparent differences between Army training and civilian education, there are numerous lessons which many educators can learn from Army experience.

What in essence, therefore, is the contribution of Army training to civilian education? The many specific training aids and devices which the Army has developed are not the truly crucial contributions. What is generally significant is that Army training has served as a laboratory for many progressive principles and practices which heretofore in education have been given either lip-service only or applied sparingly. Army training has not really discovered new principles, but has indicated how effective a job can be done by the application of the best existing principles of educational philosophy and psychology.

Army training has shown the value of:

1) realistic aims and goals for courses of study,
2) clearly defined objectives for specific subjects,
3) small-size instructional groups,
4) functional materials and texts in motivating learning,
5) correlated teaching of related subjects,
6) diversified teaching methods, i.e. lecture, demonstration, conference, coach and pupil, etc.,
7) applicatory exercises to fix and maintain desirable knowledges and skills,
8) appropriate visual and training aids,
9) tests for prediction, evaluation, and diagnosis,
10) proper classification of individuals to permit assignment in accord with needs and interests,
11) continuous psychological study of men to effect reclassification and reassignment, when necessary,
12) cumulative records to indicate each man's development in the Army,
13) national teaching conferences to insure uniformity in teaching practices,
14) 'in-service' training for teachers; i.e. troop schools.

Great opportunities exist for extending these educational gains. To take advantage of opportunities that prevail, educators should develop the same zeal about their mission which Army instructors have exhibited in their work.

In the Army, it is firmly believed that good training is essential to victory. Consequently, nothing is spared which will enhance training and make it most efficient.

Educators will have to develop a comparable faith in the value of education. They must become vitally interested in seeing that schools turn out balanced and stable individuals. They must strive to prepare individuals who are well-equipped and ready to meet the vital problems of reconstruction. And they must be willing to eliminate from education those factors and conditions which produce graduates who are ignorant of the social realities of the times. When educators are ready to do these things, they will find it easy to insist on the consistent application of those same principles and practices which have made Army training so successful.

THE RELATIONSHIP BETWEEN THE CENTROID AND SPEARMAN'S METHODS

KARL J. HOLZINGER

The University of Chicago

The purpose of the present article is to point out the relationship between the centroid method for obtaining factor coefficients and Spearman's 1914 method. The relationship between the centroid method and Spearman's theorem on the correlation of sums will also be shown. Although these relationships have been shown in some form in other papers, correspondence with workers in factor analysis has convinced the writer that they are not generally known, partly because of the forbidding symbolism in which the centroid is usually described. In the present discussion, therefore, the relationships will be indicated in the simplest manner and with a minimum of statistical mathematics.

THE CENTROID METHOD

In all methods of factor analysis a set of linear relationships between tests and factors is assumed. For four tests and one common factor these equations may be written in the form,

$$
(1)\quad \begin{aligned}
z'_1 &= a_1F + b_1U_1 \\
z'_2 &= a_2F + b_2U_2 \\
z'_3 &= a_3F + b_3U_3 \\
z'_4 &= a_4F + b_4U_4
\end{aligned}
\qquad \text{Factor Pattern}
$$

where the z's, F, and U's are variables in standard form (unit standard deviation) for convenience. The variables U_1, U_2, U_3, and U_4 are unique factors assumed to be uncorrelated with themselves and F. The a's and b's are coefficients to be determined by factorial methods. A prime is used on the z's in order to distinguish a variable in a theoretical factor pattern from an observed variable z. In describing the centroid method for obtaining the a's only three variables will be used for simplicity.

Since all factors are uncorrelated the correlations from pattern (1) (reproduced correlations) are obtained as follows:

$$r'_{12} = \frac{\Sigma z'_1 z'_2}{N} = a_1 a_2 \left(\frac{\Sigma F^2}{N}\right) = a_1 a_2$$

The expression in the parenthesis is unity because of the assumption that all variables are in standard form. The correlations r'_{11}, r'_{22}, and r'_{33} might be taken as each equal to unity, but most factor analysts prefer to use 'communalities' or values from the common-factor portion of pattern (1). A variable such as z''_1

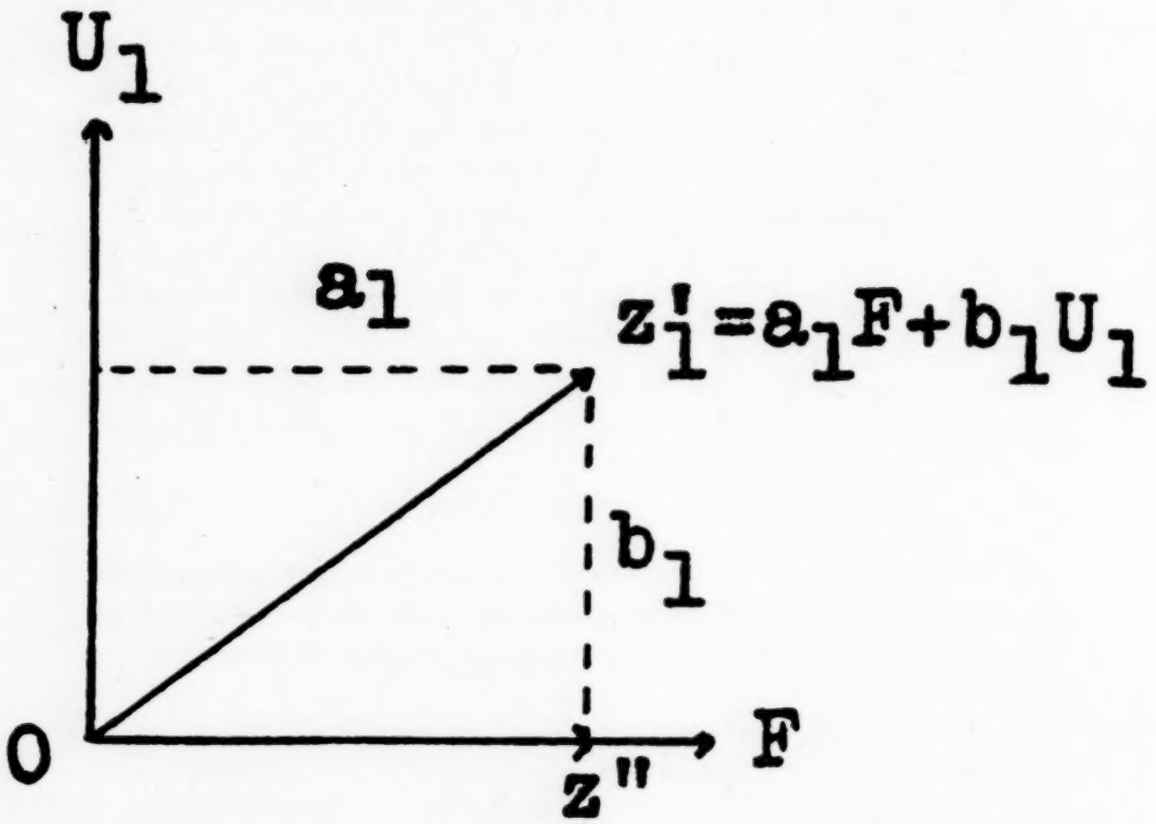

FIG. 1.

indicates the projection of z'_1 on the axis *OF* as shown in Fig. 1. The common-factor portion of pattern (1) may then be written as follows:

$$(2) \qquad \begin{matrix} z''_1 = a_1F \\ z''_2 = a_2F \\ z''_3 = a_1F \end{matrix} \quad \text{Common-Factor Portion of Pattern (1)}$$

wherein F is in standard form, but z'' now has a standard deviation equal to

$$\sqrt{\frac{\Sigma z''^2}{N}} = \sqrt{\frac{\Sigma a^2F^2}{N}} = a\sqrt{\frac{\Sigma F^2}{N}} = a = \sqrt{\text{communality}}.$$

In the reproduced correlations shown in Table 1, communalities are inserted in the diagonal of the correlation matrix.

TABLE I.—REPRODUCED CORRELATION MATRIX

Variable	1	2	3	
1.....	$a_1^2 = r'_{11}$	$a_1a_2 = r'_{12}$	$a_1a_3 = r'_{13}$	
2.....	$a_2a_1 = r'_{21}$	$a_2^2 = r'_{22}$	$a_2a_3 = r'_{23}$	
3.....	$a_3a_1 = r'_{31}$	$a_3a_2 = r'_{32}$	$a_3^2 = r'_{33}$	
S_j = Sum	$a_1(a_1+a_2+a_3)$ $= r'_{11}+r'_{21}+r'_{31}$	$a_2(a_1+a_2+a_3)$ $= r'_{12}+r'_{22}+r'_{32}$	$a_3(a_1+a_2+a_3)$ $= r'_{13}+r'_{23}+r'_{33}$	$(a_1+a_2+a_3)^2 = T$ $\Sigma(9 \text{ correlations}) = T$

The column sums of the reproduced correlations are indicated as well as the grand total, T, of the entire correlation matrix. The centroid coefficients are then given by the formula,

$$a_j = \frac{S_j}{\sqrt{T}} \tag{3}$$

For example,

$$a_1 = \frac{S_1}{\sqrt{T}} = \frac{r'_{11} + r'_{21} + r'_{31}}{\sqrt{T}} \tag{3}'$$

With actual data where a single common factor may be justified, observed correlations replace the reproduced correlations in the calculation of a_j. Centroid coefficients may of course be obtained if several common factors are involved, but the present discussion is confined to the case of a single common factor.

SPEARMAN'S 1914 METHOD

In case a pattern of the form (1) is justified, Spearman has shown that the coefficients a_j may be obtained from the reproduced correlations as follows:

$$r'_{12} = a_1a_2;\ r'_{13} = a_1a_3;\ r'_{23} = a_2a_3$$

Therefore, for example,

$$a_1^2 = \frac{(a_1a_2)(a_1a_3)}{a_2a_3} = \frac{r'_{12}r'_{13}}{r'_{23}} \tag{4}$$

Extended to four variables this formula may be written as

$$(5) \qquad a_1^2 = \frac{r'_{12}r'_{13} + r'_{12}r'_{14} + r'_{13}r'_{14}}{r'_{23} + r'_{24} + r'_{34}}$$

It will be observed that no diagonal elements or communalities are involved in the right-hand members of these last two formulas. In fact the formulas compute these communalities.

The value of a_1 from formula (4) becomes identical with the centroid value from formula (3)′ if communalities are inserted in the diagonals and if the formula is extended to include these entries:

$$(6) \quad a_1^2 = \frac{\begin{array}{r} r_{11}r_{11} + r_{12}r_{12} + r_{13}r_{13} + r_{11}r_{12} + r_{11}r_{12} + r_{11}r_{13} \\ + r_{11}r_{13} + r_{12}r_{13} + r_{12}r_{13} \end{array}}{r_{11} + r_{22} + r_{33} + r_{12} + r_{12} + r_{13} + r_{13} + r_{23} + r_{23}}$$

where the primes are dropped for simplicity. The numerator of (6) may be written as $(r_{11} + r_{12} + r_{13})^2$ and the denominator is the T of formula (3)′. The equivalence of the factor loadings computed from formulas (3)′ and (6) has thus been shown.

SPEARMAN'S CORRELATION OF SUMS

The correlation of a variable x_a with the composite $x_b + x_c + x_d$ may be written in the form,

$$(7) \qquad r_{x_a(x_b+x_c+x_d)} = \frac{\Sigma x_a(x_b + x_c + x_d)}{\sqrt{\Sigma x_a^2 \Sigma(x_b + x_c + x_d)^2}}$$

by Spearman's theorem on correlation of sums, where the x's are deviates from means.

From equations (2) it is apparent that in the common-factor space $F = \frac{z''_1 + z''_2 + z''_3}{a_1 + a_2 + a_3}$. The correlation between z'_1 and F, the composite $z''_1 + z''_2 + z''_3$, is desired. It should be noted that the variables z'' are deviates from means but are not in standard form. Spearman's formula in this case becomes

$$(8) \qquad r_{z_1'(z_1''+z_2''+z_3'')} = \frac{\Sigma z'_1(z''_1 + z''_2 + z''_3)}{\sqrt{\Sigma(z'_1)^2 \Sigma(z''_1 + z''_2 + z''_3)^2}}$$

From equations (1) and (2) it will be noted that $\Sigma z'_1 z''_1 = Na_1^2$, $\Sigma z'_1 z''_2 = Na_1a_2$, etc.; $\Sigma(z'_1)^2 = N$, $\Sigma(z''_1)^2 = Na_1^2$, and $\Sigma z''_1 z''_2$

$= Na_1a_2$. Making these substitutions in equation (8) yields

$$r_{z_1'(z_1''+z_2''+z_3'')} = \frac{N(a_1^2 + a_1a_2 + a_1a_3)}{\sqrt{N\{N[a_1^2 + a_2^2 + a_3^2 + 2(a_1a_2 + a_1a_3 + a_2a_3)]\}}}$$

$$= \frac{a_1(a_1 + a_2 + a_3)}{\sqrt{(a_1 + a_2 + a_3)^2}} = \frac{S_1}{\sqrt{T}}, \text{ as in (3)}'.$$

The centroid coefficients a_j may thus be regarded as correlations between a variable z_j and the total (or average) of the variables z_1, z_2, z_3 projected on the common-factor space. Formula (8) may be considered a definition of the centroid coefficient a_1.

The above three methods have been shown equivalent in case the Spearman formulas are extended to include the communalities in the diagonal of the correlation matrix.

THE SCORING OF CONTINUITY OR REARRANGEMENT TESTS

C. W. ODELL

University of Illinois

A number of years ago several articles concerning the scoring of the continuity or rearrangement form of test appeared, their authors proposing various methods of so doing. At the time the present writer made a comparative study of ten such proposals and published results[1] indicating that, if both validity of scores and ease of scoring are considered, several methods appear to be of somewhere near equal merit, but also suggesting the need of further study of the question.

Since the publication of the article just mentioned no truly new method of scoring such tests has come to the writer's attention. Rosander[2] further developed a previously proposed method. His procedure is simple, but requires the use of rather lengthy tables to provide for all possibilities. It and the ρ (rho) method of rank correlation appear to be about equally valid, and to be preferable to the others suggested. The complete computation of ρ and its transmutation into a score for each set of elements is, however, too expensive of time to be practicable for regular classroom use. The present writer has, therefore, devised and presents herewith a table which may be used to shorten the process of scoring by the ρ and the R methods.

This table assumes that, as is common in short-answer tests, each element should count one point and, therefore, each set of elements as many points as it contains elements. Thus, for example, the maximum score on a set of five elements is 5; on one of six, it is 6; and so on. Agreement as to this practice is quite general.

There are at least three theories as to what the minimum score should be. One is that for the worst possible answer, which is a complete reversal of correct order, the score should be 0 and that for chance order it should be half of the maximum. A

---

[1] Odell, C. W., "Still More About Scoring Rearrangement or Continuity Tests," *School Review*, Vol. 39, p. 542–46. September, 1931.

[2] Rosander, A. C., "A Simple Method of Scoring and Interpreting Sequential Responses," *Journal of Educational Research*, Vol. 36, p. 168–77. November, 1942.

second theory considers that the score which ought to be given an answer in which the elements are in chance order, thus exhibiting no correlation with their correct arrangement, is 0, and that for a complete reversal, or perfect negative correlation, it should be negative and as large as the number of elements. This is based upon the argument that a purely chance order indicates no knowledge, whereas a complete reversal is evidence of the possession of misinformation, and therefore should be penalized. The third theory likewise assumes that chance arrangement should be scored 0, but rejects the use of negative scores and uses 0 for all cases in which there is negative correlation. The second assumption appears best theoretically, but there are some objections to the use of negative scores. The present writer does not feel a strong conviction in favor of any one of these three methods as better than the other two, and therefore presents a table that may be employed for any one of them.

To employ this table, one must perform the same first operations as for rank correlation. The tables take the place of the final ones. Thus, for the method based upon R one must find the sum of the gains, or positive differences, in rank between the pupil's order and the correct order; and for the ρ method one must find the sum of the squares of all the differences, both positive and negative. With either of these sums and the number of elements in the set one enters the correct section of the table and finds therein the score (to the nearest integer) which one should give.

To illustrate this, the following example may be used: Let us assume that in a history test which calls for numbering the ele-

	Pupil's Response	Correct Order	Difference	Square of Difference
Caesar's conquest of Gaul	*4*	*4*	*0*	*0*
Battle of Actium	*5*	*5*	*0*	*0*
War between Marius and Sulla	*1*	*2*	*−1*	*1*
Formation of first triumvirate	*2*	*3*	*−1*	*1*
Death of Antony and Cleopatra	*6*	*6*	*0*	*0*
Death of the Gracchi	*3*	*1*	*+2*	*4*
			$\Sigma g = +2$	$\Sigma D^2 = 6$

Scoring Table for Continuity Tests

Part A. Scores Based upon Sum of Gains															Part B. Scores Based upon Sum of Squares of Differences						
	Minimum Score = $-N$							Minimum Score = 0								Minimum Score = $-N$			Minimum Score = 0		
Σg*	Number of Elements							Number of Elements							ΣD^2**	No. of Elems.			No. of Elems.		
	4	5	6	7	8	9	10	4	5	6	7	8	9	10		4	5	6	4	5	6
0	4	5	6	7	8	9	10	4	5	6	7	8	9	10	0	4	5	6	4	5	6
1	2	4	5	6	7	8	9	3	4	5	7	8	9	10	2	3	5	6	4	5	6
2	1	3	4	5	6	8	9	2	4	5	6	7	8	9	4	2	4	5	3	5	6
3	−1	1	3	4	6	7	8	1	3	4	6	7	8	9	6	2	4	5	3	4	5
4	−4	0	2	4	5	6	8	0	3	4	5	6	8	9	8	1	3	5	2	4	5
5		−3	1	3	4	6	7		1	3	5	6	7	8	10	0	3	4	2	4	5
6		−5	0	2	3	5	6		0	3	4	6	7	8	12	−1	2	4	2	4	5
7			−2	1	3	4	6			2	4	5	7	8	14	−2	2	4	1	3	5
8			−4	0	2	4	5			1	4	5	6	8	16	−2	1	3	1	3	5
9			−6	−2	1	3	5			0	3	5	6	7	18	−3	1	3	0	3	4
10				−4	0	2	4				2	4	6	7	20	−4	0	3	0	3	4
11				−5	−1	2	3				1	4	5	7	22		−1	2		2	4
12				−7	−2	1	3				0	3	5	6	24		−1	2		2	4
13					−4	0	2					2	5	6	26		−2	2		2	4
14					−5	−1	2					1	4	6	28		−2	1		2	4
15					−7	−2	1					1	3	5	30		−3	1		1	3
16					−8	−4	0					0	3	5	32		−3	1		1	3
17						−5	−1						2	5	34		−4	0		1	3
18						−6	−2						1	4	36		−4	0		1	3
19						−8	−3						1	3	38		−5	−1		0	3
20						−9	−4						0	3	40		−5	−1		0	3
21							−5							2	42			−1			2
22							−6							2	44			−2			2
23							−8							1	46			−2			2
24							−9							0	48			−2			2
25							−10							0	50			−3			2
															52			−3			2
															54			−3			1
															56			−4			1
															58			−4			1
															60			−4			1
															62			−5			1
															64			−5			1
															66			−5			0
															68			−6			0
															70			−6			0

* Σg is the sum of gains or increases when the correct order and that of a pupil are compared.

** ΣD^2 is the sum of the squares of the differences between the correct order and that of a pupil. It is always an even number.

ments of each set in their chronological order, one pupil's response was as shown by the figures immediately after the elements.

The scorer should have, for each set of elements, a key with the correct order so spaced as to fit alongside the pupil's response. Placing it there, he can readily, without needing to use pencil and paper, determine the sum of the positive differences. In the example it is 2. If he desires, he can check this by finding the sum of the negative differences, which should be the same. Having found this, he looks up the tabular entry corresponding to it, the number of elements, and the preferred method of scoring. If he believes the minimum score possible should be negative and the same as the number of elements, he employs the first half of Part A of the table, where in the column for six elements and the row for $\Sigma g = 2$ he finds 4 as the score to give in this instance. If he believes that 0 should be given for chance order and all orders worse than that, he also uses the same section of the table but substitutes 0 for each negative entry. In this case the score is the same, 4, as by the previous method. If he believes that 0 should be the score for complete reversal and half of the number of elements that for chance order, he uses the right half of the table. In this instance so doing yields a score of 5.

If the scorer is willing to take the additional step of squaring all the differences and summing these squares, thus employing the better but longer ρ method, he makes similar use of Part B of the table. For the example above ΣD^2 is 6. For this and six elements the score according to either of the first methods is 5, and likewise for the third method it is 5.

A second example may further clarify the use of the table. It is based upon a set of five planets to be numbered in order of

	Pupil's Response	Correct Order	Difference	Square of Difference
Mars	*2*	*4*	*−2*	*4*
Venus	*1*	*2*	*−1*	*1*
Jupiter	*3*	*5*	*−2*	*4*
Earth	*5*	*3*	*+2*	*4*
Mercury	*4*	*1*	*+3*	*9*
			$\Sigma g = 5$	$\Sigma D^2 = 22$

their distance from the sun. The sum of positive differences is 5 and the number of elements also five, whence the score is

−3, 0, or 1, according to which of the three theories is applied. The sum of the squared differences is 22, whence the score by the ρ method is −1, 0, or 2.

By the method just outlined the labor of scoring is not prohibitive. After a teacher decides which minimum score to use he can ignore the other set of columns. The same is true as to which method of correlation is taken as the basis. The writer recommends that based upon ρ, but if a teacher is unwilling to go through the increased labor required, the other, based upon R, may be used with good results. If the teacher employs groups all containing the same number of items the same column of the table can be employed throughout, which renders the work of scoring easier than if two or more must be used. The writer recommends that each teacher prepare for his own use a small card containing only the portion of Table X which is needed for the length of group he wishes to employ and the score he believes should be the minimum. Such a card can be used more quickly than the complete table, since it contains many more rows and columns than are needed.

RELATIONS BETWEEN ABILITY AND SOCIAL STATUS IN A MIDWESTERN COMMUNITY. I: TEN-YEAR-OLD CHILDREN

ROBERT J. HAVIGHURST AND LEOTA LONG JANKE*

Committee on Human Development, University of Chicago

This paper reports results from a program of testing all the children of a given age in a typical midwestern community. The purpose of the study was to find out how differences of ability are related to differences of social status. In the present paper, the results are given for ten-year-olds. A later paper will report results on sixteen-year-olds. In still another paper we propose to analyze the changes with age so as to draw conclusions about the educative effects of the different social environments in which children live in the Middle West.

The tests were selected to cover a variety of aptitudes and abilities. Tests employed were: Stanford-Binet, Cornell-Coxe, Iowa Silent Reading, Minnesota Paper Form Board, Minnesota Mechanical Assembly, Chicago Assembly Test for Girls, Porteus Maze, and Goodenough Draw-a-Man.

NATURE OF THE COMMUNITY

The community, which we shall call Midwest, is a typical middlewestern community. Its history of about a hundred years has seen the changes of agriculture and industry characteristic of the Middle West. It consists of a county seat and its surrounding rural trading territory. The total population is about 10,000, of which about 6,000 lived in the city of Midwest at the time of the 1940 census. Ninety per cent of the population are native-born. There are no Orientals and only one family of Negroes. There are two distinguishable ethnic groups; a small number of people of Polish descent, and a somewhat larger number of Norwegians.

Midwest is a relatively stable community, with an autonomous social and business life. It is not a satellite of a larger city. There is no college or state institution in Midwest, nor is it a summer or health resort.

---

* Mrs. Janke is now Instructor in Psychology at the University of Illinois.

This is the kind of community which is found most frequently in the middle western part of the United States. According to census data, there are 215 small cities with population between 5,000 and 10,000 in the twelve states of the Middle West. Of this group of cities, 108 are the largest cities in counties which are agricultural-industrial, i.e., which have between twenty-five and fifty per cent of their male workers employed in agriculture. Midwest is one of these cities. These communities, having some modern industries as well as agriculture, tend to 'have everything' offered in American life, so that individuals raised in them find it easy to step into much more complex metropolitan worlds and act with vigor and decision. A young person in Midwest learns to deal with the multiple choices and situations of a rather complex but highly integrated community. If the community were smaller and simpler, his experience would be limited by the undifferentiated nature of the rural village. On the other hand, if the community were a metropolis, his experience would be constricted by the segmentation and compartmentalization of life that is characteristic of the great city.

Midwest and the other communities of its class are a reservoir of human life. They produce more than enough children to maintain their population. Many of these children move to the larger cities, which do not reproduce themselves.

SOCIAL STRUCTURE OF THE COMMUNITY

The people of Midwest are members of a social hierarchy. They all place themselves on a social scale, above some people and below others. They speak of 'the best people,' 'the common man,' 'the chronic reliefers,' and so on. These social differences go along with differences in the rearing of children. Through our tests we see the correlates of these social differences as they operate on children.

There are two general methods of investigating the structure of the social hierarchy. The 'socio-economic' method is to study the people with respect to their occupations, income, educational level, size of library, condition of house, possession of automobile, telephone, and other material goods. By this method the people of a community can be placed on a scale of socio-economic status. The 'social status' method is to study people with respect to their social participation—who associates with whom?—and what

social groups are placed at the top, middle, and bottom of the social scale by the people in the community? By this method the people can be placed on a scale of social status. The two methods result in placing people largely in the same groups, though there are some important differences.

We used the social status method, following the procedures of Warner.[1] Professor Warner and his students studied the community and placed the families of the ten-year-old children on a scale of social status. They found five distinct social classes in Midwest. Some of the characteristics of those classes are summarized as follows:

Group A.—This group, comprising not more than two per cent of the population, is the so-called 'top crowd.' It is composed of wealthy landed-families who have resided in the community for several generations, and of the owners or managers of the larger industries, some of whom are newcomers to Midwest. These people live in the best houses, located in one section of town. The families of this group intermarry a good deal. They also have a good many social affiliations in Chicago. This group has few children, and there was none in our ten-year-old group.

Group B.—This group is composed mostly of the families of professional men, officials in industry, and leading business men. Members of this group are frequently referred to as 'community leaders,' and they participate actively as leaders in churches, Rotary, Chamber of Commerce, etc. They have better than average incomes, live in large houses, and most of them have had a good deal of formal education. About six per cent of the population fall into this class. They have relatively few children, and it happens that only one child in the ten-year-old group comes from such a family. This is less than usual as there are a number of children of families in the A and B groups, aged nine or eleven.

Group C.—In this group are found the small business men, the lesser professional men, most of the other 'white-collar' workers, and some skilled workers. They make up the bulk of members of various organizations—political, fraternal, religious and social. Notices about the bridge clubs, bunco parties, and 'get-togethers' of these people fill the society columns of the local newspaper.

[1] Warner, W. L. and Lunt, Paul S., *The Social Life of a Modern Community*. New Haven: Yale University Press, 1941.

Together with Group D they are spoken of as 'the common man.' Few of them have more than a high-school education. Approximately thirty-seven per cent of the population fall into this group. Of the ten-year-old children, twenty-three per cent come from these families.

Group D.—This is the largest of the five social classes. It comprises the people with small incomes—mainly semi-skilled workers and laborers. Though poor, these people are hard-working and respectable. They live in the poorer sections of town, in small houses. They belong to the churches, but not to many other organizations. Approximately forty-three per cent of the people of Midwest are in this class. They have relatively large families, and sixty-two per cent of the ten-year-olds come from these families.

Group E.—Members of this group have a poor reputation in the community. They are said to be 'dirty,' 'shiftless' and 'dishonest.' People above them on the social scale refer to them as 'those river rats' and 'the bottom of the heap.' They think of themselves, however, as 'just as good as anybody else.' Some of the oldest families of Midwest are in this group. Of the total population they comprise twelve per cent, and they contribute fourteen per cent of the ten-year-olds. The people in this group have the lowest income, the least formal education, and live in dilapidated shacks on the fringes of town.

TESTING PROCEDURES

All the children were tested individually by one of us (L.L.J.). The group to be tested consisted of all the children in the city and its surrounding rural territory who were born in the year 1932. Birth-dates were secured from the school records, and checked with the records of birth-certificates, and with parents in doubtful cases.

There were 114 children in this age-group whose families were residing in Midwest on November 1, 1942. One child was in a school for the deaf and dumb in another state, and was not tested. With this exception all of the children were tested. However, two children were tested only with the Cornell-Coxe, and a third child was tested only with the Cornell-Coxe and the Stanford Binet. The remaining 110 children took all six tests. A brief description of the tests which were used follows:

1) The Revised Stanford-Binet, Form L. This test was used as a measure of standard intelligence of the middle-class variety.

2) The Cornell-Coxe Performance Ability Scale. A performance test was used so as to get at the less verbal aspects of intelligence. This test was chosen for the comparability of its items with those of the Performance Scale of the Wechsler-Bellevue Test of Adult Intelligence, which was used with the sixteen-year-old group.

3) The Iowa Silent Reading Test, New Edition, Elementary Form Am. This test was chosen to test reading ability. An RQ (reading quotient) was secured by dividing the age equivalent achieved on the test by chronological age. RQ's for two children, who were too far retarded in school to be ready for this test, were obtained from the Progressive Achievement Test, Primary Battery, Form A, Grades 1-3. Sections 1 and 2 were used.

4) The Minnesota Paper Form Board, Revised, Form AA. This test was chosen as a means of measuring ability to perceive spatial relations. High scores in this test are said by its authors to be predictive of success in engineering and mechanical occupations.

5a) The Minnesota Mechanical Assembly Test. (Modified) This test was found by its originators to give a very good prediction of mechanical ability. For the boys we used our own modification of the test. We added four easy items at the beginning of the test: nut and bolt, hinge, key on ring, and turnbuckle. These were followed by the long form of the Minnesota test from which items which have become obsolete, and items which seemed to depend to a considerable extent on strength of fingers or wrists, were omitted. The following items were retained:

Box A.—Expansion nut, hose pinch clamp, linked chain, bottle stopper, push button doorbell, bicycle bell, Corbin rim lock.

Box B.—Safety razor, monkey wrench, ringstand clamp, test-tube holder, spark plug, electric plug and wire, handle for iron.

Box C.—Die holder, pliers, electric-light socket, wing nut, glass drawer knob, rope coupler, kettle cover knob, lock nut, petcock, hose clamp, pencil sharpener, air gauge valve.

5b) The Chicago Assembly Test for Girls. The only test of mechanical ability for girls, the IER Assembly Test,[2] appeared to be inadequate for our purposes. Consequently we devised a test designed to measure the skill of a girl in working with her hands. This test consists of twenty short tasks:*

(1) Stringing beads, (2) Assembling a cabinet door knob, (3) Assembling a simple pencil sharpener, (4) Assembling a kettle cover knob, (5) Attaching electric wire to plug, (6) Identifying sewing tools, (7) Identifying textile materials, (8) Stick printing, (9) Clay modeling, (10) Weaving, (11) Cutting along a line with scissors, (12) Cross stitich, (13) Outline stitch, (14) Pinking, (15) Basting, (16) Hemming, (17) Bias binding, (18) Embroidering a design on toweling, (19) Drawing a flower, (20) Tracing a design.

6) The Porteus Maze Test. This test is said by its author to be a test of foresight or of approach to a problem. Since it does not require language, it is useful as a non-verbal test of intelligence.

7) The Goodenough Draw-a-Man Test. This test is said to measure children's ability to observe and to synthesize their observations. It also is a non-verbal test. It was given as a group test to the children in the fifth grade of the public school in Midwest. Thus the entire ten-year-old group did not take the test. The test was given in the spring of 1943, when a few of the group had passed their eleventh birthday. Since Goodenough does not recommend the test for use beyond the age of ten, this fact should be remembered in interpreting the results.

RESULTS

The results of the testing program are summarized in Table I. This table shows the means, standard deviations, and ranges of test scores for the total group, and for boys, girls, urban, rural, and social class groups C, D, and E. Results on the Stanford

---

[2] Toops, H. A. *Tests for Vocational Guidance of Children Thirteen to Sixteen.* Contributions to Education, Teachers College, Columbia University, No. 136. 1923.

* Items from the Minnesota Mechanical Assembly Test (Nos. 2, 3, 4, 5) were timed. The remainder of the test was given without time limits. The average duration of the test was eighty minutes.

TABLE I.—TEST RESULTS FOR THE TEN-YEAR-OLD GROUP

	Stanford Binet				Cornell-Coxe				Goodenough Draw-a-Man				Iowa Silent Reading			
	Mean IQ	SD	N.	Range	Mean IQ	SD	N.	Range	Mean IQ	SD	N.	Range	Mean RQ	SD	N.	Range
Total Group	109	18.3	111	61–150	110	20.5	113	58–169	102	19.0	70	66–139	98	13.0	110	57–135
Boys	110	19.7	51	61–150	112	22.2	52	58–169	99	20.2	29	66–139	97	13.7	50	57–130
Girls	107	16.7	60	68–142	108	21.0	61	65–144	104	15.4	41	73–131	99	9.1	60	73–135
Urban	110	18.2	84	61–150	113	19.0	83	58–169					100	12.9	83	57–135
Rural	106	17.9	27	68–142	104	20.9	30	65–144					95	12.7	27	75–131
Status Groups																
C	114	15.0	26	84–141	116	17.9	26	87–169	107	19.0	14	78–132	99	13.0	26	82–135
D	110	16.1	68	76–142	110	19.2	70	64–155	102	16.0	46	66–134	99	11.7	68	73–131
E	91	20.2	16	61–127	96	25.3	15	58–137	91	16.0	10	71–112	88	13.2	15	57–113

	Minnesota Paper Form Board				Minnesota Mechanical Assembly (Modified)				Chicago Assembly Test for Girls				Porteus Maze			
	Mean Raw Score	SD	N.	Range	Mean t-score	SD	N.	Range	Mean t-score	SD	N.	Range	Mean test age	SD	N.	Range
Total Group	20.6	9.8	110	−1–45									12.5	2.6	110	5–17.5
Boys	22.7	8.5	49	1–41	50.0	10.0	50	28–69					12.5	2.6	51	6.5–17.5
Girls	18.9	10.4	61	−1–45					49.8	10.0	60	33–79	12.4	2.6	58	5–17.5
Urban	20.6	10.1	82	−1–42	48.9	9.9	38	28–66	50.2	9.6	45	33–72	12.5	2.6	83	6.5–17.5
Rural	20.7	9.6	28	5–45	53.7	9.2	12	39–69	48.9	11.5	15	35–79	12.4	2.7	26	5–17.5
Status Groups																
C	22.5	9.2	26	8–42	52.5	8.4	14	34–69	56.0	10.2	12	42–74	12.7	2.4	25	8–17.5
D	21.3	9.4	68	−1–45	49.2	9.6	27	28–67	49.5	9.3	41	33–79	12.8	2.2	68	7.5–17.5
E	15.7	11.6	15	1–41	46.9	11.5	8	30–61	41.3	7.1	7	35–51	10.4	3.1	15	5–16.5

TABLE II.—DIFFERENCES BETWEEN THE GROUPS—CRITICAL RATIOS

Groups Compared	Stanford Binet	Cornell-Coxe	Iowa Reading	Paper Form Board	Minn. Assembly	Chi. Assembly	Porteus Maze	Draw-a-man
Boys and Girls	.9	1.0	.7	2.1			—-	1.1
Urban and Rural	1.3	2.0	1.8	—-	1.6	—-	—-	
Status Groups								
C and D	1.1	1.4	—-	.6	1.1	2.0	—-	.9
C and E	4.0	2.7	2.4	1.9	1.2	3.8	2.6	2.2
D and E	3.5	2.0	2.7	1.7	.5	2.7	2.9	2.0

TABLE III.—COEFFICIENTS OF CORRELATION BETWEEN TESTS

	Stanford Binet	Cornell-Coxe	Iowa Silent	Paper Form Board	Mech. Assembly	Chicago Assembly
Cornell-Coxe	.72 ± .03					
Iowa Silent	.68 ± .03	.64 ± .03				
Paper Form Board	.70 ± .03	.72 ± .03	.44 ± .05			
Mech. Assembly (boys)	.66 ± .05	.76 ± .04	.21 ± .09	.66 ± .05		
Chicago Assembly (girls)	.56 ± .06	.72 ± .04	.15 ± .06	.54 ± .06		
Draw-a-Man	.50 ± .06	.63 ± .05		.48 ± .06		
Porteus Maze	.64 ± .04	.65 ± .04	.36 ± .06	.59 ± .04	.44 ± .08	.45 ± .07

Binet and Cornell-Coxe tests are given in IQ's. The Iowa Silent Reading Test results are given in RQ's, which are analogous to IQ's. The Minnesota Paper Form Board results are given in terms of raw scores. The Porteus Maze results are given in terms of the usual age-units. Scores on the two assembly tests were turned into t-scores. The reason for this was that there were no norms for the assembly tests in the forms used by us and some form of standard score would make the results most meaningful. The t-score, with a mean of 50 and a standard deviation of ten t-units, makes the results of these tests easier to compare and interpret.

The differences between the means of the various groups in the various tests have been studied by the usual method of calculating the ratio of the difference of the means to the standard error of this difference. These ratios are given in Table II. All ratios which are near zero have been omitted.

Table III presents the correlations between the results from the various tests. The product-moment coefficient was calculated in all cases.

DISCUSSION OF RESULTS

The high status of this group on the Stanford Binet and the Cornell-Coxe is a notable feature. If this is characteristic of such communities as Midwest we must think of these communities as a reservoir of high quality human resources. Achievement in reading as measured by the Iowa Silent Reading Test is not as high as the high IQ's might lead one to expect.

The relation of ability to social status is consistently what we should expect on the hypothesis that high family social position corresponds to high ability. The differences between the two lower groups, the D and E groups, are especially striking. There appears to be a group of about one-eighth of the families of the community whose children as a group are definitely low on all the abilities measured by our tests. However, even in this low group, the range is wide enough to include several children with superior ability.

Our findings on the relation of ability to social status are in substantial agreement with the findings of studies of the relation of ability to socio-economic status. However, it is doubtful that the studies of socio-economic status have differentiated off as

clearly the group at the bottom of the social ladder as has been done in our study, using social class position as the decisive factor.

The results showing the relation of ability to social status among ten-year-olds gain significance when compared with results for sixteen-year-olds. We may anticipate the publication of the results on sixteen-year-olds by saying that the simple relationships existing for ten-year-olds are changed into more complex relationships for sixteen-year-olds. There is no longer a positive correlation of all kinds of ability with social status.

The urban-rural differences are consistently in favor of urban children on the two intelligence tests and the reading test. However, the rural boys are superior to the urban boys on the test of mechanical ability. While the difference in this case is not great enough to be statistically reliable, it amounts to one-half a standard deviation which is enough to be statistically reliable in a group twice as large as this one.

The differences between urban and rural boys on the mechanical assembly test may reflect a difference in environment. The farm environment may give a ten-year-old boy more experience in the handling of tools and mechanical devices. Paterson and his co-workers[3] regarded the amount of mechanical work the boy does as a factor definitely related to performance on the mechanical assembly test. The superiority of upper status ten-year-old boys to lower status boys on the mechanical assembly test may also be related to the factor of experience with tools and mechanical devices. Children from lower status families may possess very few tools and mechanical toys.

The rural girls do not excel the urban girls on the Girls' Assembly Test. It is possible that the tasks in the girls' test do not figure more frequently in the life of a farm girl than in the life of a city girl.

The findings on urban-rural differences are consistent with those of Shephard,[4] who reports that rural children tend to be superior to urban children on mechanical ability, while urban children are superior on tests of verbal intelligence.

---

[3] Paterson, D. G. *et al, The Minnesota Mechanical Ability Tests* Ch. 13. Minneapolis: University of Minnesota Press, 1930.

[4] Shephard, E. L. "Measurement of Certain Non-Verbal Abilities of Urban and Rural Children." J. Ed. Psych., 33, 458–462, 1942.

The urban-rural differences are not social class differences. Nearly all of the children of class E are urban, while the ratio of D to C children is almost the same in the urban and rural groups. Consequently the urban children would show a greater superiority over rural children than they do in Table I if we were to eliminate the factor of social class. We may conclude that two sets of environmental factors are correlated with the test results—one set consisting of those which show urban-rural differences and the other those which show social class differences.

The only reliable sex difference is that between boys and girls on the Minnesota Paper Form Board Test. Boys have generally been found to be superior to girls on this test.

The correlations between tests are generally high. The Cornell-Coxe Performance Test results correlate more highly with the Assembly Test results for girls and for boys than do the Stanford-Binet results. This was expected.

The correlation between the modified Minnesota Assembly Test and the Stanfod-Binet (.66) was unexpected, in view of the fact that Paterson *et al*[5] report a correlation of .06 between the Otis IQ and the Minnesota Assembly Test. This difference may be in part due to the difference in age of the groups studied. Paterson's subjects were thirteen and fourteen years of age, whereas ours were ten and eleven. We have found the correlation between Stanford-Binet IQ and scores on the Minnesota Mechanical Assembly test to be much lower for sixteen-year-old boys than for our ten-year-old group.

In general we are led by these correlation results to the view that ten-year-olds show rather less differentiation and specialization of ability than do adolescents and adults. Even in the case of the most different abilities tested in our battery, mechanical ability and reading ability, there is a significant positive correlation.

SUMMARY AND CONCLUSIONS

The following tests were given to all ten-year-old boys and girls residing in a typical middle-western community, both urban and rural: Stanford Binet, Cornell-Coxe, Iowa Silent Reading, Minnesota Paper Form Board, Minnesota Mechanical Assembly

---

[5] Paterson, D. G. et al. op. cit. p. 301 and Table 89.

(Modified), Chicago Assembly Test for Girls, Porteus Maze, and Goodenough Draw-a-Man (urban children only). The test results were compared for social class groups, urban-rural and sex groups. Product-moment correlations were calculated for the various pairs of tests.

1) The mean IQ is about 110, which is much higher than the national average.

2) Children of higher family social status tended to do better in all of the tests than children of lower social position. Children of the lowest social group were definitely lower in all the abilities measured.

3) Urban children tended to do better than rural children, except on the mechanical assembly test for boys, in which the rural boys were superior.

4) In spite of the group differences, there was so much overlapping of scores that some of the lower-class children and some of the rural children were definitely superior in ability.

5) Correlation coefficients between the tests were high, and correlations between mechanical ability tests and intelligence tests were higher than have been found in previous studies. This is probably due to the fact that our subjects were younger than those of previous studies.

DIFFICULTY OF VOCABULARY LEARNING AS AFFECTED BY PARTS OF SPEECH

C. L. MORGAN
Purdue University

AND

D. N. BONHAM
Loudenville, Ohio, High School

Parts of speech affect word form and word meaning. The problem of this study was to evaluate parts of speech as related to difficulty of association of word meanings.

Krueger[7] found that different arrangements of letters forming syllables made for differences in learning difficulty. Davis[3] found that although nonsense syllables were more homogeneous, this fact was far outweighed by the fact that words were easier to memmorize. He also found that a much larger number of words may be learned in a given amount of time. Wiley[11] in his study of first-grade children found that the meaning which a word has for a first-grade child is a factor which greatly influences difficulty in learning to recognize the word symbol. Chapman,[2] in her study to determine whether a foreign language vocabulary is learned more readily when the English equivalents of the foreign language are known, concluded that foreign words are more readily learned when the English equivalent is familiar to the learner. Holmes[6] found that syllable length does not seem to be a determining factor in immediate recall or recognition of words. Calhoun[1] states that the number of one-syllable words reproduced after a short interval of time was largely independent of the length of the lists auditorily presented. He concludes further that this independence is consistently present throughout the age range which includes grades I to XII. Skinner[9] concludes that vocabulary difficulty should be a criterion for determining the order in which texts are read. The study of Gilbert[5] indicated that spellers possibly fail to distinguish between hard and easy sections of words, frequently expending as much time on one part as another.

It appears from these studies that the most important factor determining word difficulty is word meaning. The importance

of analyzing the nature of word meanings as related to vocabulary learning and usage is clear.

The problem of this study was to determine the extent to which words, as affected by their parts of speech, differ in learning difficulty. The parts of speech studied were: nouns, pronouns, verbs, adjectives, adverbs, prepositions, articles, conjunctions, and interjections. Two different methods were employed to verify the results obtained.

One hundred forty-eight seventh- and eighth-grade pupils constituted the sample of the population studied. All pupils in the sample learned all of the words.

An effort was made to equate the different parts of speech on the basis of frequency of English usage, unfamiliarity of foreign equivalent, length of foreign equivalent, ease of pronunciation and spelling, frequency of exposure sequence, and purity of grammatical form.

Among problems encountered in the selection of words were: (1) the selection of a foreign language not familiar to any student, (2) the selection of words which could be interpreted only as of one part of speech. For example, the word 'bed' may be used as a noun or as a verb. This plurality of usage was encountered often in dictionaries. The selection of English words having more than one use as parts of speech was in a few instances necessary. For example the word 'rabbit' was used in the experiment. 'Rabbit' may be a noun or it may be a verb meaning to follow another object, etc. It was decided that such interpretation would not be given by the student in his learning process. (3) The authors attempted as much as possible to use both English and foreign words of essentially the same syllable or letter length. (4) Some words are naturally harder to learn because of unfamiliarity of word meaning. Using such words as 'boy' and 'notwithstanding' might affect the study because of the more frequent use of the word 'boy.' To avoid this factor as far as possible, Thorndike's *Teachers Word Book of* 20,000 *Words*[10] was the source used for the selection of the English words. Fifty each of nouns, verbs, pronouns, adjectives, and adverbs were selected. To increase the number of prepositions, articles, conjunctions, and interjections, a number of these parts of speech were selected from Long's *A College Grammar.*[8] These English words did not conform to the standards by which other parts of

speech were selected from *Webster's Dictionary*, because of the small number of words used as these parts of speech. Prepositions, articles, and conjunctions were grouped into a single category when the parts of speech were compared as to relative difficulty. Prepositions, conjunctions, etc., that had other grammatical uses, had to be accepted in order to have enough words as a sample. To secure a sufficient number of interjections very little selection was possible. Finally fifty nouns, fifty pronouns, fifty verbs, fifty adjectives, fifty adverbs, forty-five prepositions, articles and conjunctions, and thirty interjections, totaling three hundred twenty-five words were selected. To insure the interpretation of the interjection as an interjection in the student's mind exclamation marks were used on the flash cards. To insure equal unfamiliarity with the foreign words on the part of the students, the foreign equivalents, with some change in spelling, were taken from the International Language called Ru Ro.[4]

For each learning exercise, twenty foreign words and their meanings were selected at random from the Master List; i.e., three nouns, three pronouns, three verbs, three adjectives, three adverbs, three prepositions, and two interjections.

The exposure apparatus was similar to Guhin's card changer.[12] Two methods were used in exposing the stimulus words. Method A (alternate exposure) presented first the foreign word followed by a second card containing its English equivalent. Method B (simultaneous exposure) consisted of cards containing both the foreign word and the English equivalent.

During class periods, the activity period, after school, and on Saturdays the study was carried on. On an average, the time required for one section of the students to learn all the words was about one week of two periods per day. This time fluctuated because of assembly programs, vacations, etc. The entire study required sixteen weeks. The exposure apparatus was placed on top of the teacher's desk to insure adequate visibility. The cards were then exposed at equal intervals of time, approximately 1.5 seconds for each foreign word and its English meaning. Immediately after exposing all twenty or forty flash cards, depending on whether the cards were mounted by method B or A, respectively, stenciled tests listing the foreign words just shown were given to each student by assistants. The students

were then instructed to write all the English meanings they recalled. All tests were immediately collected, the flash cards shuffled, mounted, and exposed again. Another stenciled test with the same foreign words rearranged was given to the students. Approximately six to ten exposures of a learning exercise; i.e., a group of twenty words, was possible for each class period. This exposure-testing procedure was continued until all students could correctly recall the English equivalents of all twenty words three times in succession. The number of times for each group of words to be shown before all students learned all twenty words varied from four to fifty-six times. The learning-testing procedure was not completed by all students at one sitting, nor was the time between successive intervals of study constant, due to change of classes, assemblies, etc. However, the study was continued at the earliest possible time. Provisions were made for the slower students to complete their group of words during the activity period or outside school hours. Prizes in the form of pencils, erasers, fountain pens, tablets, scrapbooks, and free theatre passes were used to motivate individual learning.

Frequency tables were made for each word and the median number of showings required to learn each word was computed. It will be recalled that all words in a single learning exercise were arbitrarily considered learned when the student gave the English meaning successfully on three consecutive tests. From the median number of exposures required to learn each word, frequency tables were made for each part of speech. The means, standard deviations, and standard error of the means of the median exposures required for learning, are separately given for each part of speech for the alternate exposure method (A) and the simultaneous exposure method (B). From Table 1 it will be observed that the two methods of presenting flash cards, Method A and Method B, are compared. The differences for the different parts of speech consistently favored the method of simultaneous exposure; but, except for consistency, the differences for adverbs, pronouns, verbs, and adjectives were unreliable. The difference for prepositions was at the one-per-cent level of significance, while the differences for nouns and interjections were at the two-per-cent level.

Table 2 indicates the relative amount of difficulty of each part of speech compared with all other parts of speech when the words

were presented by alternate exposure. It will be observed that when nouns are compared with other parts of speech the differences favor nouns in all instances at or above the five-per-cent level of significance.

TABLE 1.—COMPARISON OF ALTERNATE AND SIMULTANEOUS EXPOSURE METHODS

Part of Speech	Obtained Difference of Means	Standard Error of Difference of Means	Critical Ratio	Level of Significance	In Favor of
Prepositions	1.47	.40	3.68	.1	B**
Nouns	.94	.39	2.41	2−	B
Interjections	1.04	.44	2.36	2−	B
Adverbs	.76	.54	1.41	*	B
Pronouns	.57	.42	1.36	*	B
Verbs	.33	.45	.73	*	B
Adjectives	.21	.46	.46	*	B

* Unreliable.

** Simultaneous exposure method.

In the remaining comparisons of Table 2 no statistically reliable differences were obtained. The critical ratios found, although not statistically significant, favor pronouns over adverbs, prepositions, and verbs; verbs over adverbs and prepositions; adjectives over adverbs (at the two-per-cent level), prepositions, and pronouns; interjections over pronouns (at the two-per-cent level), verbs, adverbs, and prepositions; prepositions over adverbs. Throughout Table 2 the critical ratios are consistently unfavorable to adverbs. Since the number of adverbs found in Thorndike's *Word Book* was limited, the adverbs were longer in syllable and letter length than other parts of speech. This fact may account for the learning difference observed. Since the measures are in reality correlated and the formula for the standard error of the difference used was that employed for uncorrelated measures, the results here presented are conservative.

Table 3 indicates the relative amount of difficulty of each part of speech compared with all other parts of speech when the words were presented by simultaneous exposure. For the most part, the results of Table 3 are consistent with the findings of Table 2.

TABLE 2.—COMPARISONS WITH OTHER PARTS BY ALTERNATE EXPOSURE METHOD

Part of Speech	Observed Difference	Standard Error of the Difference	Critical Ratio	Level of Significance	In Favor of
A. Nouns with:					
Adverbs	2.33	.58	4.02	.1−	Nouns
Prepositions	1.91	.50	3.82	.1−	Nouns
Verbs	1.54	.42	3.67	.1−	Nouns
Pronouns	1.28	.40	3.20	.5	Nouns
Adjectives	1.00	.42	2.38	2	Nouns
Interjections	1.01	.50	2.02	5−	Nouns
B. Pronouns with:					
Adverbs	1.05	.50	1.78	*	Pronouns
Prepositions	.63	.51	1.24	*	Pronouns
Interjections	.27	.52	.52	*	Pronouns
Adjectives	.28	.54	.52	*	Pronouns
Verbs	.26	.54	.48	*	Pronouns
C. Verbs with:					
Adverbs	.79	.60	1.31	*	Verbs
Interjections	.53	.53	1.00	*	Interjections
Adjectives	.54	.55	.98	*	Adjectives
Prepositions	.37	.52	.71	*	Verbs
D. Adjectives with:					
Adverbs	1.33	.60	2.21	2+	Adjectives
Prepositions	.91	.52	1.74	*	Adjectives
Interjections	.01	.53	.02		Even
E. Adverbs with:					
Interjections	1.32	.59	2.28	2+	Interjections
Prepositions	.42	.58	.72	*	Prepositions
F. Prepositions with:					
Interjections	.90	.50	1.00	*	Interjections

* Unreliable.

This is to be expected in light of the results of Table 1. When nouns are compared with other parts of speech the chances favor nouns in all cases. Statistically significant differences in ease of learning favor: (1) nouns as compared with adverbs, verbs, pronouns, adjectives and prepositions at or above the one-per-cent level of significance; (2) pronouns compared with adverbs and interjections at the two-per-cent level; (3) verbs compared with interjections and prepositions at the one-per-cent level or above; (4) interjections compared with adverbs at the one-per-cent level; (5) prepositions compared with adverbs at the one-per-cent level.

The fact that parts of speech used as interjections are short words, and also that the exclamation marks following each interjection may have been clues for learning, leads one to use considerable caution when interpreting the relative difficulty of parts of speech compared with interjections. The critical ratios of Table 3 pertaining to various comparisons with adverbs, although not in all instances reliable, are again consistently unfavorable to adverbs. The remaining differences found were not considered to be reliable differences since they were below the five-per-cent level of significance, several falling below the twenty per-cent level, but the differences obtained favor pronouns over adverbs, verbs, and adjectives; verbs over adverbs; adjectives over adverbs; prepositions over verbs, adjectives, adverbs and pronouns; interjections over verbs, adjectives, adverbs, and prepositions.

SUMMARY

Presenting foreign words and their English equivalents in simultaneous or alternate exposures, where the total exposure time is constant, does not materially affect the number of exposures required for learning. Within the limitations of the evidence, the advantages favor simultaneous exposure.

To the extent that the samples are representative of their parts of speech, nouns were found easier to learn than adverbs, verbs, pronouns, prepositions, adjectives and probably interjections.

Adverbs would appear to provide the greatest difficulties of word association.

Differences in ease of learning of verbs, pronouns, prepositions, adjectives and interjections appear to be either inconsistent with the two methods used or so small as to be unreliable.

TABLE 3.—COMPARISONS WITH OTHER PARTS OF SPEECH BY SIMULTANEOUS EXPOSURE METHOD

Part of Speech	Observed Difference	Standard Error of the Difference	Critical Ratio	Level of Significance	In Favor of
A. Nouns with:					
Adverbs	2.51	.64	3.92	.1−	Nouns
Verbs	2.15	.47	3.78	.1−	Nouns
Pronouns	1.65	.55	3.00	1−	Nouns
Adjectives	1.73	.58	2.98	1−	Nouns
Prepositions	1.38	.53	2.65	1−	Nouns
Interjections	.91	.56	1.63	*	Nouns
B. Pronouns with:					
Adverbs	.86	.35	2.46	2−	Pronouns
Interjections	.74	.33	2.26	2+	Interjections
Verbs	.50	.30	1.67	*	Pronouns
Prepositions	.27	.29	.93	*	Prepositions
Adjectives	.08	.24	.33		Pronouns
C. Verbs with:					
Interjections	1.24	.33	3.72	.1−	Interjections
Prepositions	.77	.30	2.59	1−	Prepositions
Adjectives	.42	.33	1.26	*	Adjectives
Adverbs	.36	.36	1.01	*	Verbs
D. Adjectives with:					
Interjections	.82	.46	1.77	*	Interjections
Adverbs	.78	.48	1.62	*	Adjectives
Prepositions	.35	.80	.80	*	Prepositions
E. Adverbs with:					
Interjections	1.60	.38	4.26	.1−	Interjections
Prepositions	1.13	.34	3.28	.1	Prepositions
F. Prepositions with:					
Interjections	.47	.32	1.46	*	Interjections

* Unreliable.

REFERENCES

1) S. W. Calhoun. "Influence of Syllabic Length and Rate of Auditory Presentation on Ability to Reproduce Disconnected Word Lists." *Journal of Experimental Psychology*, XVIII (1935), pp. 621–632.

2) Florence Chapman and Luther C. Gilbert. "A Study of the Influence of Familiarity with English Words upon the Learning of Their Foreign Language Equivalents." *Journal of Educational Psychology*, XXVIII (November, 1937), pp. 621–628.

3) F. C. Davis. "The Relative Reliability of Words and Nonsense Syllables as Learning Material." *Journal of Experimental Psychology*, XIII (1930), pp. 221–234.

4) Edward Power Foster. *English-Ro Dictionary*. Waverly, West Virginia: Ro Language Society, (1932).

5) Luther C. Gilbert. "Experimental Investigation of a Flash-Card Method of Teaching Spelling." *Elementary School Journal*, XXXII (January, 1932), pp. 337–351.

6) F. L. D. Holmes. "Syllabic Length, Recognition, and Immediate Recall." *Journal of Applied Psychology*, XVIII (1934), pp. 831–841.

7) William C. F. Krueger. "The Relative Difficulty of Nonsense Syllables." *Journal of Experimental Psychology*, XVII (1934), pp. 145–153.

8) Mason Long. *A College Grammar*. New York: Ronald Press, (1928).

9) L. H. Skinner. "An Analysis of the Vocabulary Difficulty of Forty-Five Italian Textbooks." *Modern Language Journal*, XX (April, 1936), pp. 411–415.

10) Edward L. Thorndike. *Teachers Word Book of* 20,000 *Words*. New York: Bureau of Publications, Teachers College, Columbia University (1931).

11) E. Will Wiley. "Difficult Words and the Beginner." *Journal of Educational Research*, XCII (April, 1928), pp. 278–289.

12) C. H. Stoelting Co. *Psychological and Physiological Apparatus*. 424 N. Homan Avenue, Chicago, Illinois, (1936), p. 91.

BOOK REVIEWS

J. McV. Hunt, Editor. *Personality and the Behavior Disorders, a Handbook Based on Experimental and Clinical Research.* Two volumes. New York; The Ronald Press, 1944. Pp. xii + 1242.

This review promises to read like a publisher's press notice in the extravagance of its praise. These two volumes represent a careful, thorough, and systematic presentation of the results of contemporary research and opinion on personality and its disorders. Each chapter is by a separate author. One feels admiration in the editor's selection of the contributors to these volumes. They have certainly worked together as a team and without exception each has carried through his assignment most thoroughly and competently. This reviewer is curious concerning the way in which the editor managed his team. The team work is superb. Each of the contributors apparently not only knew the outline of the whole book and the names of those who were to contribute to it but kept these facts in mind. There is almost no overlapping or duplication of effort. There are frequent cross-references at points where a topic may be raised, but where it receives more adequate treatment in another section of the book.

To mention the high points in the book would almost be the equivalent of duplicating the table of contents. The chapter by McKinnon on "The Structure of Personality" serves as a good introduction, classifying concepts and making definitions. This chapter also includes a satisfactory review of the findings of factor analyses of personality. The chapter by Mowrer and Kluckhorn on "A Dynamic Theory of Personality" is superb and in a way sets the keynote for the book. As these authors state, they have attempted assimilation of the contributions from psychoanalysis, cultural anthropology and the theory of learning. The chapter by R. W. White discusses "Projective Techniques" and one will find here a more popular discussion of free association methods, the Rorschach test, the Thematic Appreception test and the use of play, drawing and painting for projective purposes. There are several chapters in the book which represent authoritative statements of topics of current interest. Leon J. Saul presents the latest findings on psychoso-

matic relationships. R. R. Sears discusses recent experimentation on psychoanalytic concepts including the topics of fixation, regression, aggression, substitution, repression, and projection, bringing down to date his previous review in the *Social Science Research Council Bulletin.*[1] Kurt Lewin reviews current work on level of aspiration theory. The experimentation and discussion of this topic leaves this reviewer cold, but there is no doubt that this chapter is well done and authoritative and a springboard for future workers in this area. Liddell and Finger survey the work on experimental neuroses. The chapter by Neal Miller on "Conflict" is excellent, including, as it does, some of the studies by Miller and his students hitherto unpublished. One may consult the chapters by Sheldon for a résumé and summary of recent work done on the study of the constitutional factors in personality, or the chapter by Shock on "Physiological Factors in Behavior" with restrained yet authoritative statement on the influence of glands on personality. Margaret Ribble summarizes her recent observations on six hundred neonates. She concludes that the development of the cerebral hemispheres, particularly the frontal lobes, is determined in important measure by the degree to which the infant is mothered in its early days.

The chapter by Hunt, the editor of the volumes, is also extraordinarily fine. Hunt reviews recent studies on the deterioration in mental functioning which is found in various pathological conditions. He presents here for the first time certain conclusions from his own experimental work and his integration of the findings of others. Disputing Goldstein's theory that in pathological studies there is a loss of capacity to assume the 'abstract' attitude, Hunt believes that intelligent behavior and thinking are dependent upon both fundamental capacity and learned skills based somehow on neural structures. He sees intelligence as consisting behaviorly of "ready modifiability of response and of capacity to sustain several attitudes or tasks simultaneously, to shift readily from one to another, and to synthesize them into a single complex operation." In deterioration processes tests have shown that there is "reduced modifiability, inability in performance where attitudes or tasks must be

---

[1] R. R. Sears. "Survey of Objective Studies of Psychoanalytic Concepts." *Social Science Research Council Bulletin* #51, 1943.

sustained simultaneously, or in performances that involve shifting readily from one attitude to another." We have had considerable study of mental development by comparing the capacities of younger and older children. The study of mental functioning by comparing the capacities of normal and pathological subjects is new and fruitful and Hunt's review should stimulate considerable further research in these areas in the years to come.

Lindsley has an authoritative chapter on Electroencephalography.

Certain of these chapters strike this reviewer as having less significance, but this is surely a matter of personal outlook and point of view. Guthrie repeats his behavioristic analysis of personality. A few years ago Guthrie's position was stimulations and convincing; today it leaves many questions unanswered. The two chapters by Jones on "Rating Methods" and Maller on "Personality Tests" recall developments of the past, and actually add very little new material. Jenness has included a discussion of hypnotism. One wonders why those who experiment with hypnotism must use such trivial stimuli and suggesting when there is crying need for further light on the values of hypnotism for the psychotherapy of personality disorders. The chapter by Penrose on "Heredity" is one of the least convincing in the book. As a matter of fact, the hereditary point of view loses out in these two volumes. Wherever evidence is presented that there is a possibility that a personality trend might have an hereditary basis, the statement is qualified by pointing out that the phenomena observed could also have occurred through the influence of environmental forces. Actually Penrose apparently omitted some of the most important titles from his bibliography, particularly the study of twins by Newman, Freeman and Holzinger which had to be added by the editor in a supplementary list.

The chapters on the "Psychoneuroses" by Malamud and "Functional Psychoses" by Cameron bring out most clearly the primitive state of our knowledge about the disorders of the mind. In Cameron's chapter particularly, one may find one conclusion after another by one worker contradicted by the results of another worker. Psychiatry is still struggling with the problem of classification of pathological states and, notwithstanding the

four decades of Freud and Adolph Meyer, a large body of psychiatrists in this country are still thinking of mental disorders as originating from organic and constitutional factors rather than from functional factors growing out of experience.

A book of this kind is dated by the mere fact of its excellence. One can praise it by saying that it was remarkably up-to-date and yet this very fact means that the book is growing old from the date of publication, as new research adds new knowledge in the various fields covered by these two volumes. Like similar handbooks, it should have a life span of from ten to twenty years, but the subject-matter is so important that it would be well worth revising when it became outdated.

The first printing contains a number of typographical errors which will no doubt be corrected in later printings. This reviewer was irritated by finding a number of references in the text which did not have a corresponding listing in the bibliography at the end of the chapter.

PERCIVAL M. SYMONDS
Teachers College, Columbia University

H. H. REMMERS AND N. L. GAGE. *Educational Measurement and Evaluation.* New York: Harper & Brothers, 1943, pp. 580.

This volume presents a comprehensive treatment of evaluation in education, based on the premise that individuals differ in many respects and that the needs of society are likewise varied. The emphasis is on the guidance point of view—to provide "the methods and techniques for obtaining the facts and evaluative data necessary for valid guidance" (p. viii). Both in purpose and scope the book is more than a conventional treatise on tests and measurements. The physical, the emotional, the social, the attitudinal, and the environmental phases of education receive equally prominent treatment along with the traditional topics of mental ability and academic achievement.

Educational Measurement and Evaluation poses two questions: (1) "Why Evaluate?" and (2) "What Should be Evaluated?" The major portion of the book then centers around the topic "How to Evaluate."

Chapter I deals with the purposes of evaluation, primarily as a guidance function, "to enable the right pupils to receive the right education from the right teachers" (p. 1). All other purposes of evaluation are, by implication at least, considered to be of a

secondary nature. The next six chapters deal with the formulation of educational objectives. Growth and development relative to the physical, the social, the personal, and the environmental aspects of instruction are given equal prominence with the academic and the mental aspects. The last four hundred forty pages of the book deal with the techniques and methodologies of evaluation—the construction or selection of the evaluating device, its administration, the interpretation of the data obtained, and finally the evaluation of the device itself.

The book is well adapted to the needs of those teachers and prospective teachers who are desirous of getting an orientation in the field of educational evaluation. It could very well be used as a basic text in an introductory course in evaluation. The book presents an excellent sampling of the recent literature in the field. There is a minimum of technical terminology. Statistical considerations are relegated to the last two chapters of the book.

In spite of the wealth of good material in *Educational Measurement and Evaluation* it has, nevertheless, two limitations, both due to the ambitious scope of the book. Its very comprehensiveness necessitates a superficial, and perhaps even misleading, treatment of some topics. The most serious limitation is an inadequate treatment of the factors affecting test validity and test reliability. For example, the authors state: "The emphasis of the present discussion on accuracy of measurement as the crux of the concept of reliability follows that of Jackson and Ferguson" (p. 202). Yet the thesis of these two gentlemen is not given, except as a bibliographical reference. Again, on page 206, optimum administration time is linked with test reliability; but the effect of administration time on test validity is not discussed. A discussion of item analysis is conspicuously absent. It is indeed unfortunate that the authors saw fit to devote only nineteen pages to the topic of test validity, and only nine pages to the topic of test reliability in a book of five hundred eighty pages; but it is most unfortunate that this limited presentation is so given that its readers are apt to accept as 'gospel truth' many of the exaggerated statements about tests found in manuals and in research reports.

The second limitation of the book—a confusion of terms—is academic. This limitation does not affect its worthwhileness

as an introductory text, but is one which concerns the more advanced reader. There are instances of meticulous definition and differentiation of terms followed, however, by a loose use of these same terms in subsequent discussions. The most glaring example is the very careful distinction between 'measurement' and 'evaluation' (pp. 29–30), but a synonymous use of these same terms throughout the book. On page 208 the 'estimated true score' is called a 'true score,' thus giving rise to a confused statement: viz., "One method of expressing the reliability of scores on a test which has the advantage of being independent of the range of talent used in determining reliability is the standard error of a true score. This statistic tells the range within which scores on the same test would be expected to fall two-thirds of the time if a very large number of the tests, equivalent in all respects, were given to the pupil. Thus, if a pupil's true score on a test is 70 and the standard error of this true score is 4, then his true score would be expected to range from 66 to 74 two-thirds of the time if additional forms of the same test were given to him."

GUSTAV J. FROEHLICH

University of Chicago

LAWRENCE A. AVERILL. *Introductory Psychology.* New York. The Macmillan Company, 1943, pp. 564.

Many teachers have come to believe that a course in practical psychology should be made available to students at the senior-high-school stage. The appearance of this well-written text should serve as a substantial aid in testing the soundness of this judgment. The adolescent age should be a favorable time for the development of insight into sound principles and attitudes underlying the desirable social adjustments so profoundly needed in modern living.

The psychological content is dependable and well selected for its purpose, though to a few it may appear that the author draws somewhat too heavily from an era as early as that of William James. However, the presentation is such as appears likely to secure a lively and a genuine interest on the part of the rank and file of students of high-school age. Moreover, the author has shown unusual care and skill in facilitating an insight into the working of practical psychological principles, as displayed in the lives of well-known people. This is done by the use

of interesting sketches from biography, fiction, sociological events, and literature in general. Integration of an unusual and highly desirable sort is exemplified, especially in the introduction to each chapter, and in the interesting and seasoned exercises to be found at the conclusion of each topic considered. In general this text appears to be worthy of extensive trial in the hands of qualified teachers who are willing to do pioneering work in this promising area.

BENJAMIN R. SIMPSON
School of Education
Western Reserve University